



QY 61 SSTVSTFTTFOISTPTSPGDLGFLAFPLAPYDTVIPNSAGNLLGLFPNLAIRNS 120  
 Db 61 DSAVLTSFDTIINFESTPTSTRADGLAFFIAPDSVI--SYHGGFLGLFPNANTLNN 118  
 QY 121 TTSKETTIDVNAASNNVAVFEDTYPNDNIGDPYRKHHIGIDVNSIRSKATVAMWQNGKT 180  
 Db 119 STSENQT-TTKAASNNVAVFEDTYPNDNIGDPYRKHHIGIDVNSIRSKATVAMWQNGKI 177  
 QY 181 ATAHISNSASKRLSVTTFYPPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234  
 Db 178 ATAHISNSASKRLSVTSYYAGSKPATLSYDIELHTVLPWVRVGLSASTGDK 231

## RESULT 2

US-08-881-189B-23  
 ; Sequence 23, Application US/08881189B  
 ; Patent No. 6310195

## GENERAL INFORMATION:

; APPLICANT: Colucci et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED  
 ; TITLE OF INVENTION: PROGENITOR CELL PRESERVATION FACTOR  
 ; NUMBER OF SEQUENCES: 24  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hoffmann & Baron, LLP  
 ; STREET: 350 Jericho Turnpike  
 ; CITY: Jericho  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 11753  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
 ; COMPUTER: IBM compatible  
 ; OPERATING SYSTEM: MS-DOS  
 ; SOFTWARE: WordPerfect  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/881,189B  
 ; FILING DATE: June 24, 1997  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:

## FILING DATE:

; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Feit, Irving N.  
 ; REGISTRATION NUMBER: 28,601  
 ; REFERENCE/DOCKET NUMBER: 381-44 PCT

## TELECOMMUNICATION INFORMATION:

; TELEPHONE: (516) 822-3550  
 ; TELEFAX: (516) 822-3582  
 ; INFORMATION FOR SEQ ID NO: 23:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 286 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; US-08-881-189B-23

Query Match 65.9%; Score 796.5; DB 4; Length 286;

Best Local Similarity 66.7%; Pred. No. 3.2e-78;

Matches 156; Conservative 30; Mismatches 45; Indels 3; Gaps 2;

QY 1 AQSVSFTFKFSDOKDLMFQGHITSSNVQLTKDSNGNPVSTSVGRVLYSAPLRLWE 60  
 Db 23 AQSLSFTFKFSDOKDLMFQGHITSSNVQLTKDSNGNPVSTSVGRVLYSAPLRLWE 82  
 QY 61 SSTVSTFTTFOISTPTSPGDLGFLAFPLAPYDTVIPNSAGNLLGLFPNLAIRNS 120  
 Db 83 DSAVLTSFDTIINFESTPTSTRADGLAFFIAPDSVI--SYHGGFLGLFPNANTLNN 140  
 QY 121 TTSKETTIDVNAASNNVAVFEDTYPNDNIGDPYRKHHIGIDVNSIRSKATVAMWQNGKT 180  
 Db 141-STSENQT-TTKAASNNVAVFEDTYPNDNIGDPYRKHHIGIDVNSIRSKATVAMWQNGKI 199  
 QY 181 ATAHISNSASKRLSVTTFYPPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234

Db 200 ATAHISNSVSKRLSVTSYAGSKPATLSYDIELHTVLPWVRVGLSASTGDK 253

## RESULT 3

US-09-141-821-2  
 ; Sequence 2, Application US/09141821  
 ; Patent No. 6110891

## GENERAL INFORMATION:

; APPLICANT: Arpad Janos PUSZTAI  
 ; APPLICANT: Szuzanna Magdalena BARDOCZ  
 ; APPLICANT: Richard Michael John PALMER  
 ; APPLICANT: Neil William FISH  
 ; APPLICANT: Gyorgy J. KOTELES  
 ; TITLE OF INVENTION: LECTIN COMPOSITIONS AND USES THEREOF  
 ; FILE REFERENCE: 48656  
 ; CURRENT APPLICATION NUMBER: US/09/141,821  
 ; CURRENT FILING DATE: 1998-08-28  
 ; NUMBER OF SEQ ID NOS: 5  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 285  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 ; US-09-141-821-2

Query Match 37.8%; Score 457; DB 3; Length 285;

Best Local Similarity 44.7%; Pred. No. 2e-41;

Matches 105; Conservative 40; Mismatches 68; Indels 22; Gaps 8;

QY 3 SVSFTFTFKFSDOKDLMFQGHITSSNVQLTKDSNGNPVSTSVGRVLYSAPLRLWE 61  
 Db 34 SLSFSPFKFKHQDPLIFQSDALVTSKGVQLQTTV--NDRVYDSIGKRVLAAPFQIWD 91  
 QY 62 ST-VVSTFTTFOISTPTSPGDLGFLAFPLAPYDTVIPNSAGNLLGLFPNLAIRNS 120  
 Db 92 TTGNVASEFTVSFSLIKAPNEGKTADGLVFLAPVGSF-QPLKGGGLGLF-----KDE 144  
 QY 121 TTSKETTIDVNAASNNVAVFEDTYPNDNIGDPYRKHHIGIDVNSIRSKATVAMWQNGKT 180  
 Db 145 SYNK-----SNQIVAVEFTFRN-VAMDPNGIHMGIDVNSTQSVRTVRWDWANGEV 194  
 QY 181 ATAHISNSASKRLSVTTFYPPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234  
 Db 195 ANVFISTEASTKSITASLVYPSLEKSFILSAIVDLKKVLPWVRVGFATTGTLSE 249

## RESULT 4

US-09-141-821-1

; Sequence 1, Application US/09141821

; Patent No. 6110891

## GENERAL INFORMATION:

; APPLICANT: Arpad Janos PUSZTAI  
 ; APPLICANT: Szuzanna Magdalena BARDOCZ  
 ; APPLICANT: Richard Michael John PALMER  
 ; APPLICANT: Neil William FISH  
 ; APPLICANT: Gyorgy J. KOTELES  
 ; TITLE OF INVENTION: LECTIN COMPOSITIONS AND USES THEREOF  
 ; FILE REFERENCE: 48656  
 ; CURRENT APPLICATION NUMBER: US/09/141,821  
 ; CURRENT FILING DATE: 1998-08-28  
 ; NUMBER OF SEQ ID NOS: 5  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1  
 ; LENGTH: 285  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 ; US-09-141-821-1

Query Match 36.9%; Score 446; DB 3; Length 285;

Best Local Similarity 42.6%; Pred. No. 3.1e-40;

Matches 100; Conservative 47; Mismatches 66; Indels 22; Gaps 8;



```

RESULT 8
US-08-881-189B-12
; Sequence 12, Application US/08881189B
; Patent No. 6310195
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID
; TITLE OF INVENTION: PROGENITOR CE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLC

```

```

RESULT 9
US-09-141-821-3
; Sequence 3, Application US/09141821
; Patent No. 6110891
;
; GENERAL INFORMATION:
;
; APPLICANT: Arpad Janos PUSZTAI
; APPLICANT: Szusanna Magdolna BARDOCS
; APPLICANT: Richard Michael John PALMER
; APPLICANT: Neil William FISH
; APPLICANT: Gyorgy J. KOTLES
;
; TITLE OF INVENTION: LECTIN COMPOSITIONS AND USES THEREOF
;
; FILE REFERENCE: 48656
;
; CURRENT APPLICATION NUMBER: US/09/141.821
; CURRENT FILING DATE: 1998-08-28
;
; NUMBER OF SEQ ID NOS: 5
;
; SOFTWARE: FastSEQ for Windows Version 3.0
;
; SEQ ID NO 3
;
; LENGTH: 274
;
; TYPE: PRT
;
; ORGANISM: Human
;
US-09-141-821-3

```

```

Query Match      28.2%; Score 341; DB 3; Length 274;
Best Local Similarity 36.4%; Pred. No. 7e-29;
Matches 84; Conservative 42; Mismatches 79; Indels 26; Gaps 7;

QY 2 QSVSFTTKFSDQDKDLMFQGH-TISSNNQLTKLDNGNPVTSVGRVLYSAPLRWE 60
      . . . . . : : : : : : : : : : : : : : : : : : : : : : : :

```



Db 31 EGISFNFTNTRGQGVLLGOANIMANGILALT---NHTNP-TWNTGRALYSKVPVIWD 86  
QY 61 SST-VVSTFTFTFOISTPTSPGDLAFPLADYDVIIPNSAGNLLGLFPNLAURN 119  
Db 87 SATGNVASFVTSFVKEIGGIPADGIVFLAP-EARIIDNSAGGOLGI----- 136  
QY 120 STTSKETTIDVNAASNNVAVFDTYPNDNIGDPYRKHIGIDVNSIRSKATVADWONGK 179  
Db 137 -----VNANKAYNPFVGFEDTYSNN--WDPKSAHIGIDASSLSLRTVKWKNVSGS 186  
QY 180 TATAHISYNSASKRLSVTTFYPGKAVSLSDHVELTQVLPOWIRVGFSAST 230  
Db 187 LUKVSIYLSLKTLSVVVTHENGQISTIAQVWDLKAVLGEKVRVGFTAAT 237

## RESULT 10

US-08-881-189B-13  
; Sequence 13, Application US/08881189B  
; Patent No. 6310195  
; GENERAL INFORMATION:  
; APPLICANT: Colucci et al.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED  
; TITLE OF INVENTION: PROGENITOR CELL PRESERVATION FACTOR  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hofmann & Baron, LLP  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11753  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/881,189B  
; FILING DATE: June 24, 1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Felt, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: 381-44 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 105 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-881-189B-13

Query Match 23.6%; Score 285.5; DB 4; Length 105;  
Best Local Similarity 58.4%; Pred. No. 1.7e-23;  
Matches 66; Conservative 13; Mismatches 21; Indels 13; Gaps 4;

QY 1 AQSVSFTTFKDSQKDLMEFGHTISSNNVQLTKLDSNGNPNVSTSVGRVLYSAPRLWE 60  
Db 1 AQLSFSFTKDFPNQEDLIQGTATS-----KLSAGNPVSSSAGRVLYSAPRLWE 52  
QY 61 SSTVSTFTFTFOISTPTSPGDLAFPLADYDVIIPNSAGNLLGLFPN 113  
Db 53 DSAVLTSFDP--TIYIFTNTSRIADGLA-FIAPPDVSI--SYHGGLGLFPN 100

## RESULT 11

US-09-228-986-77

; Sequence 77, Application US/09228986  
; Patent No. 6359198  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Niels  
; TITLE OF INVENTION: Compositions Isolated from plant cells  
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling  
; FILE REFERENCE: 11000/1020  
; CURRENT APPLICATION NUMBER: US/09/228,986  
; CURRENT FILING DATE: 1999-01-12  
; NUMBER OF SEQ ID NOS: 130  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 77  
; LENGTH: 632  
; TYPE: PRT  
; ORGANISM: Pinus radiata  
US-09-228-986-77

Query Match 21.0%; Score 254; DB 4; Length 632;  
Best Local Similarity 31.8%; Pred. No. 6.9e-19;  
Matches 81; Conservative 37; Mismatches 91; Indels 46; Gaps 10;

QY 1 AOSV-----SFTFTKDSQKDLMEFGHTISSNNVQLTKLDSNGNPNVSTSVGRVLYS 53  
Db 23 AQAVERDRHDTTFLDFGNGTNLILEANASVIGSESVLSLT-----NHSHEFMLGRALYA 77  
QY 54 APLRLWESSITVSTFTFTFOISTPTSPGDLAFPLADYDVIIPNSAG--NLGLLF 111  
Db 78 APVQMKNNHT-VSSFTTFVFSIVPPPSNEGGLAFIMTPYTS---PMGAQPVQYLGL- 132  
QY 112 PNLNALRNSSTSKETTIDVNAASNNVAVFDTYPNDNIGDPYRKHIGIDVNSIRSKAT- 170  
Db 133 --NLTSN-----GQPNHLFAVEFDITMNVFEDKDPDRNHVGVDSINLSISVOTE 179  
QY 171 VAWDW-----QNGKTATAHISYNSASKRLSVT---TFYPGKAVSLSDHVELTQV 217  
Db 180 TAGYNGEEFHELNRSGRNIAQWIDYDHESSLNVITVAGLPRPQRLISLQIDLQNI 239  
QY 218 LPQWIRVGFSAATGL 232  
Db 240 VEEKMLVGFSAATGL 254

## RESULT 12

US-08-853-659A-43  
; Sequence 43, Application US/08853659A  
; Patent No. 5925522  
; GENERAL INFORMATION:  
; APPLICANT: Wong, K.K.; Saffer, J.D.  
; TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection  
; TITLE OF INVENTION: Of A  
; TITLE OF INVENTION: Salmonella Sequence, And Methods Of Detection  
; NUMBER OF SEQUENCES: 67  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Paul W. Zimmerman  
; ADDRESSEE: Intellectual Property Services  
; ADDRESSEE: Battelle Memorial Institute  
; ADDRESSEE: PNNL P.O. Box 999  
; STREET: Washington Way  
; CITY: Richland  
; STATE: Washington  
; COUNTRY: U.S.A.  
; ZIP: 99352  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage  
; COMPUTER: IBM PC/XT/AT  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Word Processor (WordPerfect 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/853,659A  
; FILING DATE: Unknown  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:

```
; APPLICATION NUMBER: none
; FILING DATE: n/a
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-853-659A-43

Query Match
Best Local Similarity 23.7%; Pred. No. 0.02; Length 252;
Matches 52; Conservative 26; Mismatches 72; Indels 69; Gaps 12;

QY 17 DLMFQGHFISSN--VIOLTKLDNGNPGVSTSGRVLYSAPRLWESSVWSTFETI--- 71
Db 25 DTGLOGDNMTSTQTFALQHIDDDAVRTVSV-----EHGGVTTTFDQTKGT 72

QY 72 --FTFQISTPYTS-PPGDLGAFFLAPYD-TVIPPNSAGNLLGLFPNLNLRNLTSTTKETT 127
Db 73 GGWTF---TPPTSWADGD-----YTLVSVEDKAGN-----TSHSASLTVT 110

QY 128 IDVNAASNNVAVFEDTYPNDNIGDPYKHKIGI-----DVSIR-----SKA 169
Db 111 VDTQTAINNIELVDSGIPDDNLTNNRPHFQVTVPTDVNVVRLSIDGKTKWFNATQSAT 170

QY 170 TVAMDW-----QNGKTATAHISYNSASKRLSVITTFP 201
Db 171 PGVDYIWPDDVADGGYLTIVTEAT-DEAGNKATQISISP 208

RESULT 13
US-09-134-001C-4463
; Sequence 4463, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4463
; LENGTH: 2137
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4463

Query Match
Best Local Similarity 7.9%; Score 95; DB 4; Length 2137;
Matches 58; Conservative 44; Mismatches 111; Indels 40; Gaps 12;

QY 1 AQSVSFETTKFSDQKDLFMFGHTISS--NVIQTLKLDNGNPGVSTSGRVLYSAPRL- 58
Db 610 SQTVIYKFKDVGQPOISVDSQTRVSKTINPITITTTDNSKDLVLTIV---TGLPSGLS 665

QY 59 --WESSTVSTFETTTFTFOISTPYTSPGDD--GLAFLAPYDVIPI-----PNSAGNLLG 109
Db 666 FDOTNTITGTPSEVGTITVTNTTATGNVTSKQPTITIQDTISFVNVNTPSOASE--- 722

QY 110 LPPNLNLRNLTSTKETTIDVNAASNNVAVFEDTYPNDNIGDPYKHKIGIDVNSIRSKA 169
Db 723 VTPINPITITADNSGKVVHTVGLPGLKFDASTNSIVGTPTQ-----IGTNTITIES 778

QY 170 TVAMDQNGKTATAHISY-----NSASKRL-----SVTTFPGKAVLSHSDVELAQVL- 218
Db 779 TDA----SGNKTTKINIEVTRNSASDSTSTSIIVNSVSTSI--SNSSTLSDSVKASQSL 832

QY 219 -PQWIRVGFSA 230
```

```
Db 833 TSKSLSESLSAST 845
; : : ||||
RESULT 14
US-08-856-253-7
; Sequence 7, Application US/08856253
; Patent No. 6288214
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Patti, Joseph M.
; APPLICANT: House-Pompeo, Karen
; APPLICANT: Sthanam, Narayana
; APPLICANT: Symersky, Jindrich
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,253
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/017,678
; FILING DATE: 16-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: TAMK:193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 345 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-856-253-7

Query Match
Best Local Similarity 7.8%; Score 94.5; DB 4; Length 345;
Matches 59; Conservative 32; Mismatches 73; Indels 81; Gaps 16;

QY 3 SVSTFTTKFSDQKDLFMFGHTISS--NVIQTLKLDNGN--PVSTSGRVL----- 51
Db 107 NVIYTFDYVNTKDDVK---AATLMPAYIDPENVKKTGNVTLATIGSTANKTVLVDYE 163

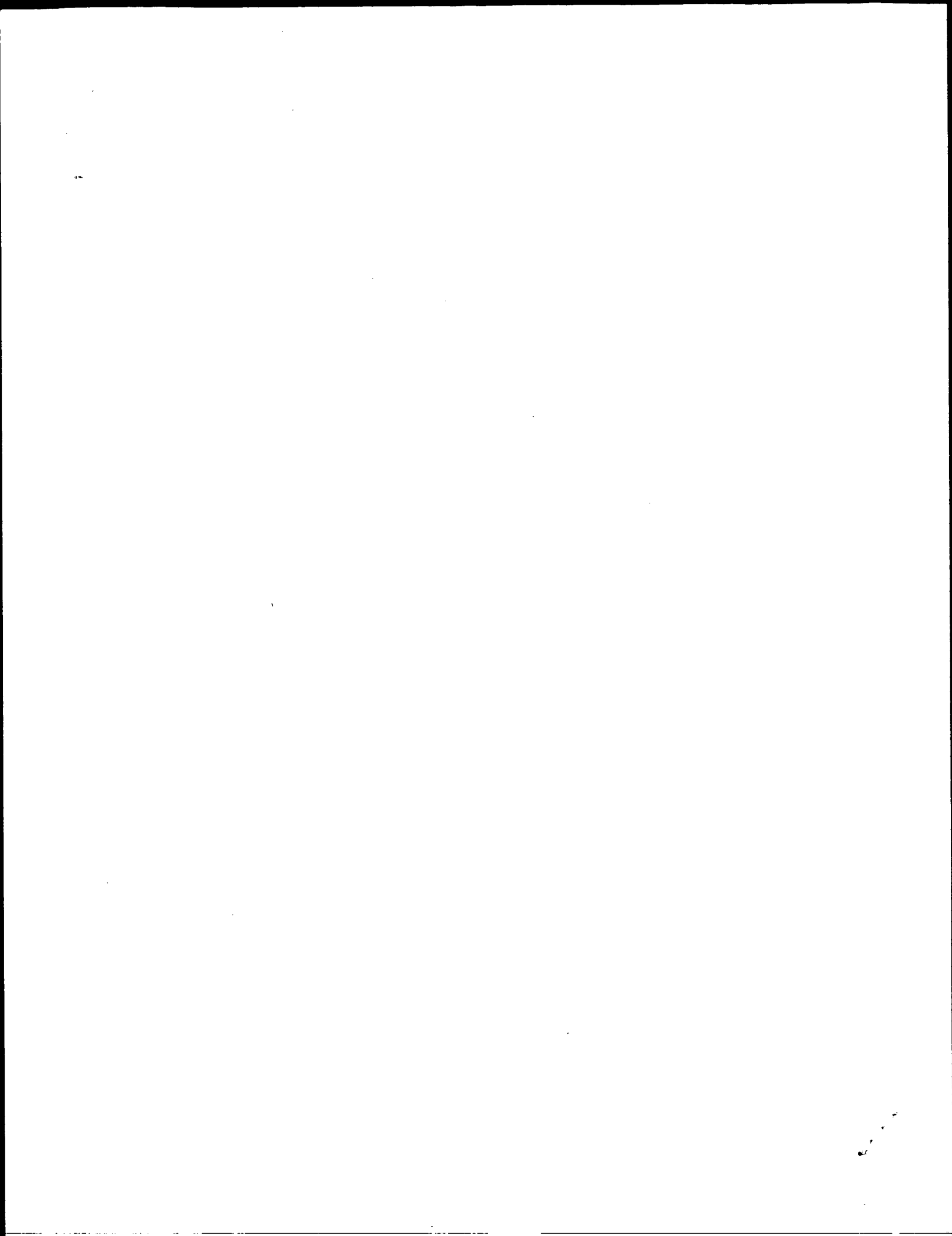
QY 52 -YSAPLRLWESSVSTVSTFETTTFTFOISTPYTSPGDLGAFFLAPYDVIIPPNSAGNLLGL 110
Db 164 KYGFYNLISIKGTIDQIDKTNNTYR-QTIYVNPSTG-----DNVIAPVLTGNLK-- 210

QY 111 FPNL--NAL--RNSTTSKETTIDVNA--SNNVAVE 141
Db 211 -PNTDSNALIDQQTNSIKVYKVD-NAADLSSESYFVNPFENFEDVTNSVNTITFPNQYKVE 268

QY 142 PDTYPNDNIGDPYKHKIG--IDVNS-----KATVAV---DWQNGKTATAHIS 186
Db 269 FNT-PDQITTPYIVVNGHIDPNKSGDLAURSTLYGYNSTIWRSMWDN-----EVA 321

QY 187 YNSAS 191
; : : :
```

Search completed: February 13 0775







A;Title: Sialic acid-binding motif of Maackia amurensis lectins.

Qy	2	QSVSFTTFKFDQDKLMEFQGH-T-ISSSNVIQTLKDSNGNPVSTSVGRVLSYASPLRLWE	60
Db	38	EALSFTEFKVSNQDELLLOGDALVSSKGELOTRVE-NGQPIPHSVGRALYSDPVHWD	96
Qy	61	SST-VVSFTFTFTTQISTPYTSPGCDGLAFFLAPYDTVIPNSAGNLUGLFPFNALRN	119
Db	97	SSTGSVASEVTSFTFVEAPENKNKTADGLAFLAPDITQV--QSLGGFLGLF-----N	147
Qy	120	SITSKETTLIDVMAASNVVAVEDFYPDNDINGPYRKHHIGIDVNSIRSKATVAMDQNGK	179
Db	148	SS-----VYNSSNQILAVEDFYFNS--WDPTARHIGIDVNSIESTRATWCWRNGE	197





QY 232 LE 233

Db 245 DE 246

## RESULT 11

A25701

phytohemagglutinin chain L precursor - kidney bean

N:Alternate names: PHA-L

C:Species: Phaseolus vulgaris (kidney bean)

C:Date: 08-Mar-1989 #sequence\_revision 08-Mar-1989 #text\_change 11-Apr-1995

C:Accession: A25701

R:Voelker, T.A.; Staswick, P.; Chrispeels, M.J.

EMBO J. 5, 3075-3082, 1986

A:Title: Molecular analysis of two phytohemagglutinin genes and their expression in Phaseolus

A:Reference number: A25701

A:Accession: A25701

A:Molecule type: DNA

A:Residues: 1-273 &lt;VOE&gt;

A:Experimental source: cv. Pinto UI111

C:Genetics:

A:Gene: dlec2

C:Superfamily: plant lectin

## Query Match

Best Local Similarity 37.0%; Score 447; DB 2; Length 273;

Matches 101; Conservative 41; Mismatches 72; Indels 24; Gaps 7;

QY 1 AQSVSFTFKFDS-DQKDLMFQGH-TISSNVQLTKLDSNGNPVSTSVGRVLYSAPLRL 58

Db 19 ANSASQTFESDFRNETLILQGDASVSSGQLRLTNVNSNGEPTVGLGRAFYSAPQI 78

QY 59 WESST-VSTFTFTFTFOISTPTSPGDLAPFLAPYDTPVIPSNSAGNLLGLFPNLNAL 117

Db 79 WDYTCNVASEDTNFENILVPNAGPADGLAFALVPVGS--QPKDKGGFLGLFDGSGN-- 134

QY 118 RNSTSKETTTIDVNAASNNVAVEFTYPNDNIGDPYRKHIGIDVNSIRSKATVAVDWQN 177

Db 135 -----SNFHTVAVEFTLYNKD-WDPREHIGIDVNSIKTIKTPWDFVN 178

QY 178 GKTATAHISYNSAKRLSVTTFPGKG-AVLSHDVELTOVLQWIRVGFSASTGLEK 234

Db 179 GENAEVHITYESSTKLLVASLVYPSLSTFTVSDTVDLKSVLPFWVSVGFSAITGITK 236

## RESULT 12

S62690

agglutinin I precursor - black locust

N:Alternate names: lectin

C:Species: Robinia pseudoacacia (black locust)

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 07-May-1999

C:Accession: S62690; S62685

R:van Damme, E.J.M.; Barre, A.; Rouge, P.; van Leuven, F.; Peumans, W.J.

Plant Mol. Biol. 29, 1197-1210, 1995

A:Title: The seed lectins of black locust (Robinia pseudoacacia) are encoded by two genes

A:Reference number: S62685; MUID:96191285; PMID:8616218

A:Accession: S62690

A:Molecule type: mRNA

A:Residues: 1-285 &lt;VAN&gt;

A:Cross-references: EMBL:U24249

A:Experimental source: seed

A:Accession: S62685

A:Molecule type: protein

A:Residues: 32-51 &lt;VAN&gt;

C:Superfamily: plant lectin

C:Keywords: glycoprotein; homotetramer; lectin

F:1-31/Domain: signal sequence #status predicted &lt;SIG&gt;

F:32-285/Product: agglutinin I #status experimental &lt;MAT&gt;

## Query Match

Best Local Similarity 36.8%; Score 445; DB 2; Length 285;

Matches 100; Conservative 41.9%; Mismatches 58; Indels 28; Gaps 1;

Matches 98; Conservative 49; Mismatches 65; Indels 22; Gaps 8;

QY 3 SVSFTFTKFDSDQKDLMFQGH-TISSNVQLTKLDSNGNPVSTSVGRVLYSAPLRLWES 61

Db 34 SLVSFPPKFAFPNQPYLIFQDALVTSTGVQLTNNV-VNGVPPRRSIGRALYAAPFQIWDN 92

QY 62 ST-VVSTFTFTFTFOISTPTSPGDLAPFLAPYDTPVIPSNSAGNLLGLFPNLNLRNS 120

Db 93 TTGNVASEFTSFIIQAPNPATTAGLAFPLAPVDT--QPGDLGGMLGIF-----KDG 144

QY 121 TTSKETTIDVNAASNNVAVEFTYPNDNIGDPYRKHIGIDVNSIRSKATVAVDWQNGKT 180

Db 145 SYNK-----SNQIVAVEFTFSNIHF-DPKGRHMGINVASIVSVKTPVNNWTNGEV 194

QY 181 ATAHSYNSAKRLSVTTFPGKGAVLSHD-VELTQVLQWIRVGFSASTGLE 233

Db 195 ANVFISYEASTKSLNASLVYPSLSTFTVSDTVDLKSVLPFWVSVGFSAITGITD 248

## RESULT 13

S62691

agglutinin II precursor - black locust

N:Alternate names: lectin

C:Species: Robinia pseudoacacia (black locust)

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 07-May-1999

C:Accession: S62691; S62686

R:van Damme, E.J.M.; Barre, A.; Rouge, P.; van Leuven, F.; Peumans, W.J.

Plant Mol. Biol. 29, 1197-1210, 1995

A:Title: The seed lectins of black locust (Robinia pseudoacacia) are encoded by two genes

A:Reference number: S62685; MUID:96191285; PMID:8616218

A:Accession: S62691

A:Molecule type: mRNA

A:Residues: 1-285 &lt;VAN&gt;

A:Cross-references: EMBL:U24250

A:Experimental source: seed

A:Accession: S62686

A:Molecule type: protein

A:Residues: 32-50 &lt;VAN&gt;

A:Superfamily: plant lectin

C:Keywords: glycoprotein; homotetramer; lectin

F:1-31/Domain: signal sequence #status predicted &lt;SIG&gt;

F:32-285/Product: agglutinin II #status experimental &lt;MAT&gt;

## Query Match

Best Local Similarity 36.7%; Score 444; DB 2; Length 285;

Matches 100; Conservative 46; Mismatches 56; Indels 22; Gaps 8;

QY 3 SVSFTFTKFDSDQKDLMFQGH-TISSNVQLTKLDSNGNPVSTSVGRVLYSAPLRLWES 61

Db 34 SLVSFPPKFAFPNQPYLIFQDALVTSTGVQLTNNV-VNGVPPRRSIGRALYAAPFQIWDN 92

QY 62 ST-VVSTFTFTFTFOISTPTSPGDLAPFLAPYDTPVIPSNSAGNLLGLFPNLNLRNS 120

Db 93 TTGNVASEFTSFIIQAPNPATTAGLAFPLAPVDT--QPLDLGGMLGIF-----KNG 144

QY 121 TTSKETTIDVNAASNNVAVEFTYPNDNIGDPYRKHIGIDVNSIRSKATVAVDWQNGKT 180

Db 145 YFNK-----SNQIVAVEFTFSNRH-WDPTGRHMGINVASIVSVKTPVNNWANGEV 194

QY 181 ATAHSYNSAKRLSVTTFPGKGAVLSHD-VELTQVLQWIRVGFSASTGLE 233

Db 195 ANVFISYEASTKSLNASLVYPSLSTFTVSDTVDLKSVLPFWVSVGFSAITGITD 248

## RESULT 14

JQ1981

lectin II - Scotch broom

N:Alternate names: 2-acetamido-2-deoxy-D-galactose binding lectin II

C:Species: Cytisus scoparius (Scotch broom)

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 07-May-1999

C:Accession: JQ1981

R:Konami, Y.; Yamamoto, K.; Osawa, T.; Irimura, T.

```

Db      183  EDAEVLITYDSSTKLLWASLAVYFSQKTSFVSDTVLKSVPWVRVGFSAATSGITK  239

Search completed: February 26, 2003, 16:51:34
Job time : 19.5281 secs

```

Query Match	36.4%;	Score 439.5;	DB 2;	Length 249;
Best Local Similarity	41.7%;	Pred. No. 3.2e-29;		
Matches 101;	Conservative 46;	Mismatches 66;	Indels 29;	Gaps 9;

QY 1 AQSVSFTTKFDSQKDLMFQGH--TISSNVIQLTKLDSNGNPVSTSVGRVLYSAPLRL 58

Db 1 SEELSFSTKFKTDQKNLLEQRDALITPTGKLQLTVE-NGKPAAYSLGRALYSTPIHI 59

QY 59 WESST-VVSTFETTFQIS--TPYTSPPGDGLAFFLAPYDTVIPPNSAGNLLGLEPNL 114

```
Db      60 WDKSTGDEASFATFFSFVISDAPNPSTAAT-DGLAFFLAPADT--QPQSAGGYLGLF--- 113
```

QV 115 NALRNSTTSKETTIDVNAASNNVAVEEDTYPNDNIGDPYRKHIGIDVNSIRSKATVAWD 174

db 114 -----EKDSSYN-----SSNOI VAVEEDTYYN SAWDPOTNPHIGIDVNTKSKKVSSWG 162

Q7 175 W0NCKTATAHISYNSASKPRI.SVTTEFVPGGK-----AVSLSHDVEITQVI.POWTRVGFESAS 229

162 EKNCNNVATWVITVVOPECKEIVASIVVPESCOTENKTSVITSANNVVKATVPEWVPICESAT 222

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

— 5 —

## RESULT 15

phytohemagglutinin L precursor - Phaseolus acutifolius

C; Species: Phaseolus acutifolius

C;Accession: S51831

Plant Mol. Biol. 26, 1103-1113, 1994

A;Reference number: S51827; MUID:95111094; PMID:7811969

A;Accession: D31051  
A;Status: nucleic acid sequence not shown

A; Molecule type: **mirna**  
A; Residues: **1-276 <MIR>**

C:Superfamily: plant lectin

F:1-24/Domains: signal sequence #status predicted <SIG>  
E:25-276/Product: phytohemagglutinin L #status predicted <MAT>

Query Match 36.3%: Score 438.5: DB 2: Length 276:

Best Local Similarity 41.4%; Pred. No. 4.4e-29;  
Matches 98. Conservative 44: Mismatches 70: Indels 25: Gaps

1 105VSEETETKEDSDOKDI MEQCH-TISSNNVIOITKI DSCNCNPVSTSVGPVI VSAPI.PIW 59

25 ANDICE NEODE - NETNIT TI ACDA SVSSECOI PI TINI NDNCEPRTI SSICPAEYSTBIAIW 82

CA BCCB - WISEEETEEETATCEETVEECCDCCCTAEEYADVDTEVTDNCACTYICIEDNTNAIDP 118

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100											

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

[illegible]

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:41:33 ; Search time 6.7191 Seconds  
(without alignments)  
1444.458 Million cell updates/sec

Title: US-09-476-485a-8

Perfect score: 1209

Sequence: 1 AQSVSFTTKFDSQKDLME.....TQVLPQWIRVGFSASTGLEK 234

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Match Length	ID	Description
1	666	55.1	240	LEC_BOWMI	p42088 bowringia m
2	646.5	53.5	290	CONA_CANGL	p14894 canavalia g
3	642.5	53.1	290	CONA_CANEN	p02866 canavalia e
4	629.5	52.1	237	LECA_DOLLA	p38662 dolichos la
5	593.5	49.1	290	LEC2_CIALU	Q39529 cladrastis
6	566	46.8	293	LEC1_CIALU	Q39528 cladrastis
7	477.5	39.5	290	LECR_CIALU	Q42372 robinia pse
8	469.5	38.8	286	LEC2_ROBPS	Q42372 robinia pse
9	468.5	38.8	240	LECS_VATMA	p81371 vatairea ma
10	468	38.7	270	LECS_SOPJA	p93538 sophora jap
11	455	37.6	292	LECS_SOPJA	p93535 sophora jap
12	447	37.0	273	PHAM_PHAVU	p15231 phaseolus v
13	446	36.9	285	LCB1_PHOBPS	Q41159 robinia pse
14	445	36.8	285	LCB1_ROBPS	Q41162 robinia pse
15	444	36.7	285	LCB2_ROBPS	Q41161 robinia pse
16	443	36.6	248	LEC2_CYTSC	p29257 cytiscus sco
17	441.5	36.5	280	LEC2_MEDTR	Q01807 medicago tr
18	438.5	36.3	243	LEC4_GRIST	p24146 griffonia s
19	437	36.1	285	LEC_SOYBN	p05046 glycine max
20	437	35.7	277	LEC1_MEDTR	Q01806 medicago tr
21	420.5	34.8	249	LEC2_ULEPU	p22973 ulex europe
22	420.5	34.8	272	PHAL_PHAVU	p05087 phaseolus v
23	418.5	34.6	275	PHAE_PHAVU	p05088 phaseolus v
24	413	34.2	250	LEC1_LABAL	p23558 laburnum al
25	411.5	34.0	275	LEC_PEA	p02867 pisum sativ
26	405.5	33.5	233	LEC_VICFA	p02871 vicia faba
27	404	33.4	281	LEC_ERYGO	p16404 erythrina c
28	402.5	33.3	236	LEC_ONOVI	p02874 onobrychis
29	402	33.3	275	LEC1_DOLBI	p05045 dolichos bi
30	379	31.3	290	LEC1_DOLBI	p16030 bauhinia pu
31	374	30.9	275	LEC3_DOLBI	p19588 dolichos bi
32	368	30.4	273	LECG_ARAHY	p02872 arachis hyp
33	366.5	30.3	242	LEC1_PSOTE	Q02431 psophocarpu

## RESULT 1

## LEC\_BOWMI

## AC

## P42088;

## DT

## 01-NOV-1995 (Rel. 32, Created)

## DT

## 15-DEC-1998 (Rel. 32, Last sequence update)

## DE

## Lectin (Agglutinin) (BMA).

## OS

## Bowringia mildbraedii.

## OC

## Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

## OC

## Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

## OC

## eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Bowringia.

## OX

## NCBI\_TaxID=28956;

## [1]

## RN

## SEQUENCE.

## RP

## TISSUE=Seed;

## RX

## MEDLINE=93385179; PubMed=8373823;

## RA

## Chawla D., Animashaun T., Hughes R.C., Harris A., Aitken A.;

## RT

## "Bowringia mildbraedii agglutinin: polypeptide composition, primary

## structure and homologies with other legume lectins.";

## RL

## Biochim. Biophys. Acta 1202:38-46(1993).

## CC

## -1- FUNCTION: BINDS PREFERENTIALLY TO OLIGOSACCHARIDES BEARING THE

## SEQUENCE MAN-ALPHA-1-&gt;2 MAN-ALPHA-1-&gt;6 MAN-ALPHA-1-&gt;6MAN FOUND

## IN EARLY STEPS OF GLYCOPROTEIN PROCESSING IN THE ENDOPLASMIC

## RETICULUM. IT BINDS WEAKLY TO HIGHLY PROCESSED OLIGOSACCHARIDE

## STRUCTURES.

## CC

## -1- COFACTOR: CALCIUM AND MANGANESE ARE ESSENTIAL FOR THE SACCHARIDE-

## BINDING AND CELL-AGGLUTININATING ACTIVITIES.

## CC

## -1- SUBUNIT: HETEROTETRAMER OF TWO ALPHA AND TWO BETA CHAINS;

## DISULFIDE BOND LINKED.

## CC

## -1- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.

## HSP; P02866; IDQ2.

## DR

## InterPro; IPR000985; Lectin\_lega.

## DR

## InterPro; IPR001220; Lectin\_legb.

## DR

## Pfam; PF00138; lectin\_lega; 1.

## DR

## Pfam; PF00139; lectin\_legb; 1.

## DR

## ProDom; PD000671; lectin\_lega; 1.

## DR

## ProDom; PD000711; lectin\_legb; 1.

## DR

## PROSITE; PS00307; LECTIN\_LEGUM\_BETA; 1.

## DR

## PROSITE; PS00308; LECTIN\_LEGUM\_ALPHA; FALSE\_NEG.

## KW

## Lectin; Calcium; Manganese.

## FT

## CHAIN 1 116

## FT

## CHAIN 117 240

## FT

## METAL 127 127

## FT

## METAL 129 129

## FT

## METAL 131 131

## FT

## METAL 133 133

## FT

## METAL 138 138

## FT

## METAL 143 143

## FT

## DISULFID 5 5

## FT

## VARIANT 143 143

## FT

## VARIANT 158 158

## FT

## VARIANT 177 177

## FT

## VARIANT 187 187

## FT

## VARIANT 203 203

## FT

## UNSURE 86 90

## 34

## 357

## 29.5

## 265

## 1

## LECN\_PEA

## 35

## 350.5

## 29.0

## 244

## 1

## LEC\_LATSP

## 36

## 344.5

## 28.5

## 181

## 1

## LECB\_LATOC

## 37

## 342

## 28.3

## 272

## 1

## LCB3\_ROBPS

## 38

## 339

## 28.0

## 236

## 1

## LECA\_CRAFL

## 39

## 330.5

## 27.3

## 240

## 1

## LEC\_LOPTE

## 40

## 330

## 27.3

## 237

## 1

## CONA\_CANBR

## 41

## 330

## 27.3

## 237

## 1

## CONA\_CANLI

## 42

## 330

## 27.3

## 237

## 1

## CONA\_CANVI

## 43

## 325.5

## 26.9

## 243

## 1

## LEC1\_ULEEU

## 44

## 324.5

## 26.8

## 211

## 1

## LEC\_LENCU

## 45

## 321.5

## 26.6

## 185

## 1

## LEC\_VICVI

## ALIGNMENTS

Wed Feb 26 17:04:16 2003

```

FT UNSURE 116 116
SQ SEQUENCE 240 AA; 25424 MW; 5A9F7FAF3A09B060 CRC64;

Query Match 55.1%; Score 666; DB 1; Length 240;
Best Local Similarity 60.1%; Pred. No. 8.6e-48;
Matches 140; Conservative 25; Mismatches 50; Indels 18; Gaps 4;

QY 1 AQSVSFTFTKFDSDQKLMFQGH-TISSNNVIOQLTKLDSNGNPVSVSGRVLYSAPLRWLW 59
Db 1 ANSVCFETFTDFESGQDLIFQGDASVGNKALQITKVDKGNFQGGVGRALYTAPLRWLW 60
QY 60 ESTVSVFTFTTFTQISTPTVTPPGDGLAFFLAPYDTVIPPNSAGNLLGLFNLNLRN 119
Db 61 OSSSLVASFEFTFTFISQGS-SSTPARALFFIASPDTKIPKSPGGGRLLGLFGSSN---- 115
QY 120 STTSKETTIDVNAASN--VVAVEFDTPNDNIGDPYKHHIGIDVNSRSKATVADWQNG 178
Db 116 -----NAGSDNGVVAVEFDTPNTDIDGPNYRHIGIDVNSRSKAAKWDWQNG 164
QY 179 KTATAHISYNSASKRLSVTFYPCGKAVSLSHDVELTQVLPQWIRVGFSASTG 231
Db 165 KTATAHISYNSASKRLSVSSYPNPPVVSFDELNNVGPDPVVRVGFSASTG 217

RESULT 2
CONA_CANGL STANDARD; PRT; 290 AA.
AC PI4894;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Concanavalin A precursor (Con A).
OS Canavalia gladiata (Sword bean) (Japanese jack bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Canavalia.
OX NCBI_TaxID=3824;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seedling;
RX MEDLINE=90127395; PubMed=2404793;
RA Yamauchi D., Minamikawa T.;
RT "Structure of the gene encoding concanavalin A from Canavalia gladiata and its expression in Escherichia coli cells.";
RL FEBS Lett. 260:127-130(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RA Yamauchi D., Nakamura K., Asahi T., Minamikawa T.;
RT "Nucleotide sequence of cDNA for concanavalin A from Canavalia gladiata seeds.";
RL Plant Cell Physiol. 30:147-150(1989).
CC -1- FUNCTION: D-MANNOSE SPECIFIC LECTIN.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- PTM: THE MATURE CHAIN CONSISTS OF RESIDUES 164-281 FOLLOWED BY POST-TRANSLATIONAL MODIFICATION AFTER REMOVAL OF THE SIGNAL SEQUENCE; CLEAVAGE AFTER ASN AT POSITIONS 148, 163, AND 281 IS FOLLOWED BY TRANSPOSITION AND LIGATION (BY FORMATION OF A NEW PEPTIDE BOND) OF RESIDUES 164-281 AND 30-148.
CC -1- MISCELLANEOUS: BINDS ONE MANGANESE (OR OTHER TRANSITION METAL) ION AND ONE CALCIUM ION. THE METAL IONS ARE ESSENTIAL FOR THE SACCHARIDE-BINDING AND CELL-AGGLUTINATING ACTIVITIES.
CC -1- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

```

```

DR EMBL: X16041; CAA34163.1; -.
DR PIR: A34139; A34139.
DR PIR: JQ2130; JQ2130.
DR HSSP: P02866; IDQ2.
DR InterPro: IPR000985; Lectin_legA.
DR InterPro: IPR001220; Lectin_legB.
DR Pfam: PF00138; lectin_legA; 1.
DR Pfam: PF00139; lectin_legB; 1.
DR ProDom: PD000671; Lectin_legA; 1.
DR ProDom: PD000711; Lectin_legB; 1.
DR PROSITE: PS00307; LECTIN_LEGUME_BETA; 1.
DR PROSITE: PS00308; LECTIN_LEGUME_ALPHA; 1.
KW Lectin; Calcium; Manganese; Glycoprotein; Signal.
FT SIGNAL 1 29
FT CHAIN 30 148 CONCAVALIN (SECOND PART).
FT PROPEP 149 163 CONCAVALIN (FIRST PART).
FT CHAIN 164 281
FT PROPEP 282 290
FT SITE 148 149 CLEAVAGE.
FT SITE 163 164 CLEAVAGE.
FT SITE 281 282 CLEAVAGE.
FT METAL 171 171 MANGANESE (BY SIMILARITY).
FT METAL 173 173 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 175 175 CALCIUM (BY SIMILARITY).
FT METAL 177 177 CALCIUM (BY SIMILARITY).
FT METAL 182 182 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 187 187 MANGANESE (BY SIMILARITY).
FT CARBOHYD 152 152 N-LINKED (GLCNAC...) (PROBABLE).
SQ SEQUENCE 290 AA; 31421 MW; 3A1C9E9ADADA3580 CRC64;

Query Match 53.5%; Score 646.5; DB 1; Length 290;
Best Local Similarity 54.0%; Pred. No. 4.4e-46;
Matches 127; Conservative 36; Mismatches 63; Indels 9; Gaps 4;

QY 3 SVSFTFTKFDSDQKLMFQGH-TISSNNVIOQLTKLDSNGNPVSVSGRVLYSAPLRWLW 61
Db 36 ALHFWNFQSKQKDLILQGDATTGTGDLNLETRVSSNGSPQSSVGRALFTAPVHWIS 95
QY 62 STVSVFTFTTFTQISTPTVTPPGDGLAFFLAPYDTVIPPNSAGNLLGLFNLNLRN 121
Db 96 SAVASFDATFTFLIKSP--DSHPADGIAEFISNIDSSIFSGSTGRLLGLFPDANVRNS- 153
QY 122 TSKETTIDVNAASN--NVVAVEFDTPNDNIGDPYKHHIGIDVNSRSKATVADWQNG 179
Db 154 ---TTIDENAAVNAADTIIVAVELDTYPNTDIDGPNYPHIGIDIKSVRSKRTAKWNQNG 209
QY 180 TATAHISYNSASKRLSVTFYPCGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234
Db 210 VGTAAHIYNSVGRRLSASVSPNGSDSATVSYDVLNDVLPVWVRVGLSASTGLYK 264

RESULT 3
CONA_CANEN STANDARD; PRT; 290 AA.
AC P02866;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Concanavalin A precursor (Con A).
OS Canavalia ensiformis (Jack bean) (Horse bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Canavalia.
OX NCBI_TaxID=3823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95086270; PubMed=3965973;
RA Carrington D.M., Auffret A., Hanke D.E.;
RT "polypeptide ligation occurs during post-translational modification of concanavalin A."
RL Nature 313:64-67(1985).
RN [2]
RP SEQUENCE OF 30-148.

```

RX MEDLINE=75095622; PubMed=1112813;  
 RA Wang J.L., Cunningham B.A., Waxdal M.J., Edelman G.M.;  
 RT "The covalent and three-dimensional structural of concanavalin A. I.  
 RL Amino acid sequence of cyanogen bromide fragments F1 and F2.";  
 RN J. Biol. Chem. 250:1490-1502(1975).  
 RX SEQUENCE OF 164-281.  
 RX MEDLINE=75095623; PubMed=1112814;  
 RA Cunningham B.A., Wang J.L., Waxdal M.J., Edelman G.M.;  
 RT "The covalent and three-dimensional structure of concanavalin A. II.  
 RL Amino acid sequence of cyanogen bromide fragment F3.";  
 RN J. Biol. Chem. 250:1503-1512(1975).  
 RX X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).  
 RX MEDLINE=73053316; PubMed=4638345;  
 RA Hardman K.D., Ainsworth C.F.;  
 RT "Structure of concanavalin A at 2.4-A resolution.";  
 RL Biochemistry 11:4910-4919(1972).  
 RX X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RX MEDLINE=75095624; PubMed=1112815;  
 RA Becker J.W., Reeke G.N. Jr., Wang J.L., Cunningham B.A., Edelman G.M.;  
 RT "The covalent and three-dimensional structure of concanavalin A. III.  
 RL Structure of the monomer and its interactions with metals and  
 RT saccharides.";  
 RL J. Biol. Chem. 250:1513-1524(1975).  
 RX X-RAY CRYSTALLOGRAPHY.  
 RX MEDLINE=75095625; PubMed=1112816;  
 RA Reeke G.N. Jr., Becker J.W., Edelman G.M.;  
 RT "The covalent and three-dimensional structure of concanavalin A. IV.  
 RL Atomic coordinates, hydrogen bonding, and quaternary structure.";  
 RL J. Biol. Chem. 250:1525-1547(1975).  
 RX X-RAY CRYSTALLOGRAPHY (1.2 ANGSTROMS).  
 RA Parkin S., Rupp B., Hope H.;  
 RT "Atomic resolution structure of concanavalin A at 120 K.";  
 RL Acta Crystallogr. D 52:1161-1168(1996).  
 RX X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).  
 RX MEDLINE=20347885; PubMed=10748006;  
 RA Bouckaert J., Dewalle Y., Poortmans F., Wyns L., Loris R.;  
 RT "The structural features of concanavalin A governing non-proline  
 RL peptide isomerization.";  
 RL J. Biol. Chem. 275:19778-19787(2000).  
 CC -!- FUNCTION: D-mannose specific lectin.  
 CC -!- SUBUNIT: Homotetramer.  
 CC -!- PTM: THE MATURE CHAIN CONSISTS OF RESIDUES 164-281 FOLLOWED BY  
 CC 30-148. TO FORM A MATURE CHAIN THE PRECURSOR UNDERGOES FURTHER  
 CC POST-TRANSLATIONAL MODIFICATION AFTER REMOVAL OF THE SIGNAL  
 CC SEQUENCE; CLEAVAGE AFTER ASN AT POSITIONS 148, 163, AND 281 IS  
 CC FOLLOWED BY TRANSPOSITION AND LIGATION (BY FORMATION OF A NEW  
 CC PEPTIDE BOND) OF RESIDUES 164-281 AND 30-148.  
 CC -!- MISCELLANEOUS: BINDS ONE MANGANESE (OR OTHER TRANSITION METAL) ION  
 CC AND ONE CALCIUM ION. THE METAL IONS ARE ESSENTIAL FOR THE  
 CC SACCHARIDE-BINDING AND CELL-AGGLUTINATING ACTIVITIES.  
 CC -!- SIMILARITY: BELONGS TO THE LEGUMINOSUS LECTIN FAMILY.  
 CC -!- DATABASE: NAME=worthington enzyme manual;  
 CC WWW="http://www.worthington-biochem.com/manual/C/CONA.html"  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed, usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X01632; CAA25787.1; .  
 DR PIR; A03357; CVJBP.  
 DR PIR; A03358; CVJBP.  
 DR PDB; 1CNI; 30-SEP-83.  
 DR PDB; 2CNA; 31-JUL-94.

DR PDB; 3CNA; 15-APR-91.  
 DR PDB; 5CNA; 15-OCT-94.  
 DR PDB; 1CON; 31-JAN-94.  
 DR PDB; 2CTV; 31-OCT-93.  
 DR PDB; 1SCR; 15-OCT-94.  
 DR PDB; 1SCS; 15-OCT-94.  
 DR PDB; 1APN; 03-APR-96.  
 DR PDB; 1CES; 15-FEB-97.  
 DR PDB; 1CJP; 15-OCT-97.  
 DR PDB; 1CVN; 14-OCT-96.  
 DR PDB; 1ENQ; 17-AUG-96.  
 DR PDB; 1ENR; 16-FEB-99.  
 DR PDB; 1ENS; 17-AUG-96.  
 DR PDB; 1GIC; 20-AUG-97.  
 DR PDB; 1JBC; 12-FEB-97.  
 DR PDB; 1NLS; 26-NOV-97.  
 DR PDB; 1ONA; 17-SEP-97.  
 DR PDB; 1TEI; 24-JUN-98.  
 DR PDB; 1VAL; 11-JAN-97.  
 DR PDB; 1VAM; 11-JAN-97.  
 DR PDB; 1VIN; 01-APR-97.  
 DR PDB; 1BXH; 07-OCT-98.  
 DR PDB; 2CAU; 25-NOV-98.  
 DR PDB; 2CAV; 25-NOV-98.  
 DR PDB; 1DQ0; 19-JAN-00.  
 DR PDB; 1DQ1; 19-JAN-00.  
 DR PDB; 1DQ2; 19-JAN-00.  
 DR PDB; 1DQ4; 19-JAN-00.  
 DR PDB; 1DQ5; 19-JAN-00.  
 DR PDB; 1DQ6; 19-JAN-00.  
 DR InterPro; IPR000985; Lectin\_legA.  
 DR InterPro; IPR001220; Lectin\_legB.  
 DR Pfam; PF00138; lectin\_legA; 1.  
 DR Pfam; PF00139; lectin\_legB; 1.  
 DR ProDom; PD000671; Lectin\_legA; 1.  
 DR ProDom; PD000711; Lectin\_legB; 1.  
 DR PROSITE; PS00307; LECTIN\_LEGUME\_BETA; 1.  
 DR PROSITE; PS00308; LECTIN\_LEGUME\_ALPHA; 1.  
 KW Lectin; Calcium; Manganese; Glycoprotein; Signal; 3D-structure.  
 FT SIGNAL 1 29  
 FT CHAIN 30 148 CONCAVALIN (SECOND PART).  
 FT PROPEP 149 163 CONCAVALIN (FIRST PART).  
 FT CHAIN 164 281  
 FT PROPEP 282 290  
 FT SITE 148 149 CLEAVAGE.  
 FT SITE 163 164 CLEAVAGE.  
 FT SITE 281 282 CLEAVAGE.  
 FT METAL 171 171 MANGANESE.  
 FT METAL 173 173 MANGANESE AND CALCIUM.  
 FT METAL 175 175 CALCIUM.  
 FT METAL 177 177 CALCIUM.  
 FT METAL 182 182 MANGANESE AND CALCIUM.  
 FT METAL 187 187 MANGANESE.  
 FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .).  
 FT CONFLICT 33 33 E -> Q (IN REF. 2).  
 FT CONFLICT 35 35 N -> D (IN REF. 2).  
 FT CONFLICT 62 62 E -> D (IN REF. 3).  
 FT CONFLICT 66 66 R -> E (IN REF. 3).  
 FT CONFLICT 98 98 V -> T (IN REF. 3).  
 FT CONFLICT 100 101 AS -> SA (IN REF. 3).  
 FT CONFLICT 107 107 T -> A (IN REF. 3).  
 FT STRAND 167 173  
 FT STRAND 178 180  
 FT HELIX 181 181  
 FT TURN 187 192  
 FT STRAND 199 202  
 FT TURN 207 208  
 FT TURN 210 218  
 FT TURN 219 222  
 FT STRAND 223 229  
 FT TURN 231 232  
 FT STRAND 236 241









Query Match 38.8%; Score 469.5; DB 1; Length 286;  
 Best Local Similarity 45.1%; Pred. No. 1.4e-31;  
 Matches 106; Conservative 42; Mismatches 66; Indels 21; Gaps 8;

QY 3 SVSFTFTKFDSDQKDLMFQGH-T-ISSNVITQLTKLDSNGNPVSTSVGRVLYSAPRLRWES 61  
 DB 34 SLSEFPKFKSQPDLIFQSDALVTSKGLVQLTV-NDGRPVYDSIGRVLYAAPPQIWD 92

QY 62 ST-VVSTFTFTFTFQISTPYTSPGDLGAFFLAPYDTVIPPNSAGNLLGLFPNLNLRNS 120  
 DB 93 TTGNVASFVTSFIIKAPNEGTADCLVFLAPVGST-QPLKGGGLGLF-----KDE 145

QY 121 TTSKETTIDVNAASNNVAVVEFDTPNDNIGDPYRKHIGIDVNSIRSKATVANDWONGKT 180  
 DB 146 SYNK-----SNOIVAVEFDTPRN-VAMDPNGIHMGIDVNSIQSVRTVRMDWANGEV 195

QY 181 ATAHISYNSAKRLSVTTFYP-GKAVLSLSDHVELTQVLPQWIRVGFSGASTGLEK 234  
 DB 196 ANVFIYSYEASTKSLTASLVPSLEKSFILSAIVDLKKVLPWVRVVGFTATTGLSE 250

RESULT 9  
 LECS\_VATNA  
 ID LECS\_VATNA STANDARD; PRT; 240 AA.  
 AC P81371;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Seed lectin (VNL).  
 OS Vatairea macrocarpa.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucotids 1; Fabales; Fabaceae; Papilionoideae; Dalbergiaceae; Vataleae.  
 OX NCBI\_TaxID=77050;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Seed;  
 RX MEDLINE=98218569; PubMed=9559667;  
 RA Calvete J.J., Santos C.F., Mann K., Grangeiro T.B., Nimtz M.,  
 Urbanke C., Sousa-Cavada B.;  
 RT "Amino acid sequence, glycan structure, and proteolytic processing of  
 the lectin of Vatairea macrocarpa seeds.";  
 RL FEBS Lett. 425:286-292(1998).  
 CC -!- FUNCTION: LECTIN THAT BINDS GALACTOSE.  
 CC -!- SUBUNIT: HOMOTETRAMER.  
 CC -!- TISSUE SPECIFICITY: SEED.  
 CC -!- PTM: PARTIALLY N-GLYCOSYLATED AT POSITIONS 111 AND 183 WITH  
 THE HEPTASACCHARIDE [(BETA-Xylosyl-1,2)(ALPHA-Mannosyl-1,6)(ALPHA-  
 Mannosyl-1,3)]BETA-Mannosyl-1,4-GlucNAc-BETA-1,4-GlucNAc-BETA-1,4  
 [ALPHA-Fucosyl-1,3]GLCNAC. A SMALL PROPORTION OF ALPHA CHAINS ARE  
 PROTEOLYTICALLY CLEAVED AT 114-115 INTO GAMMA AND BETA CHAINS.  
 CC THIS IS PROBABLY DEPENDENT ON THE DEGLYCOSYLATION OF N-111.  
 CC -!- MISCELLANEOUS: BINDS ONE MANGANESE (OR OTHER TRANSITION METAL) ION  
 AND ONE CALCIUM ION. THE METAL IONS ARE ESSENTIAL FOR THE  
 SACCHARIDE-BINDING AND CELL-AGGLUTINATING ACTIVITIES (BY  
 SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.  
 DR HSP; P19588; 1LUL.  
 DR GlycosuitedB; P81371; -.  
 DR InterPro; IPR000985; Lectin\_legA.  
 DR InterPro; IPR001220; Lectin\_legB.  
 DR Pfam; PF00138; lectin\_legA; 1.  
 DR Pfam; PF00139; lectin\_legB; 1.  
 DR ProDom; PD000671; Lectin\_legA; 1.  
 DR ProDom; PD000711; Lectin\_legB; 1.  
 DR PROSITE; PS00307; LECTIN\_LEGUME\_BETA; 1.  
 DR PROSITE; PS00308; LECTIN\_LEGUME\_ALPHA; 1.  
 KW Lectin; Calcium; Manganese; Glycoprotein.  
 FT CHAIN 1 240 SEED LECTIN ALPHA CHAIN.  
 FT CHAIN 1 114 SEED LECTIN GAMMA CHAIN.  
 FT CHAIN 115 239 SEED LECTIN BETA CHAIN.  
 FT METAL 123 123 MANGANESE (BY SIMILARITY).

FT METAL 125 125 MANGANESE AND CALCIUM (BY SIMILARITY).  
 FT METAL 129 129 CALCIUM (BY SIMILARITY).  
 FT METAL 132 132 MANGANESE AND CALCIUM (BY SIMILARITY).  
 FT METAL 137 137 MANGANESE (BY SIMILARITY).  
 FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .).  
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .).  
 FT VARIANT 117 117 I -> V.  
 FT VARIANT 148 148 M -> K.  
 FT VARIANT 154 154 G -> A.  
 FT VARIANT 168 168 E -> Q.  
 FT UNSURE 239 240  
 SQ SEQUENCE 240 AA; 26197 MW; C17DF6B2568C65C1 CRC64;

Query Match 38.8%; Score 468.5; DB 1; Length 240;  
 Best Local Similarity 42.6%; Pred. No. 1.4e-31;  
 Matches 101; Conservative 43; Mismatches 68; Indels 25; Gaps 7;

QY 1 AQSVSFTTKFDSDQKDLMFQGH-T-ISSNVITQLTKLDSNGNPVSTSVGRVLYSAPRLRW 59  
 DB 1 SEVVSFSTKFPNPNKDIILQGDALVTSKGLQTKV-KDGRKVDHSLGRALYAAPIH 59

QY 60 ESST-VVSTFTFTFQISTPYTSPGDLGAFFLAPYDTVIPPNSAGNLLGLFPNLNLR 118  
 DB 60 DSDTRVASEATSFVVEAPDESKTAGIAFFLAPDPT--QPKDGGFGLFENDSN--- 114

QY 119 NSTTSKETTIDVNAASNNVAVVEFDTPNDNIGDPYRKHIGIDVNSIRSKATVANDWONG 178  
 DB 115 -----KSLQTVAVEFDTF--SNWDPSPARHIGINVNIESMKYVKGWENG 158

QY 179 KTATAHISYNSAKRLSVTTFYP-GKAVLSLSDHVELTQVLPQWIRVGFSGASTGLEK 234  
 DB 159 KVANVYSYEASTKTLTASLVPSNATSVANSVNDLKSALPEWVRVGFSGATSGLSR 215

RESULT 10  
 LECS\_SOPJA  
 ID LECS\_SOPJA STANDARD; PRT; 270 AA.  
 AC P93538;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Bark lectin precursor (LECSJABG) (fragment).  
 OS Sophora japonica (Japanese pagoda tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucotids 1; Fabales; Fabaceae; Papilionoideae; Sophoreae; Sophora.  
 OX NCBI\_TaxID=3897;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Bark;  
 RX MEDLINE=97201486; PubMed=9049272;  
 RA van Damme E.J., Barre A., Rouge P., Peumans W.J.;  
 RT "Molecular cloning of the bark and seed lectins from the Japanese  
 pagoda tree (Sophora japonica).";  
 RL Plant Mol. Biol. 33:523-536(1997).  
 CC -!- FUNCTION: GALNAC-SPECIFIC LECTIN.  
 CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC EMBL; U63014; AAB51458.1; -.  
 DR HSP; P19588; 1LUL.  
 DR InterPro; IPR000985; Lectin\_legA.  
 DR InterPro; IPR001220; Lectin\_legB.  
 DR Pfam; PF00138; lectin\_legA; 1.  
 DR Pfam; PF00139; lectin\_legB; 1.  
 DR ProDom; PD000671; Lectin\_legA; 1.

```

DR ProDom: PD000711; Lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
DR LECTIN; Calcium; Manganese; Glycoprotein; Signal.
KW NON_TER 1
FT SIGNAL 1 15
FT CHAIN 16 270
FT METAL 141 141
FT METAL 143 143
FT METAL 145 145
FT METAL 147 147
FT METAL 150 150
FT METAL 155 155
FT METAL 155 155
FT CARBOHYD 60 60
FT CARBOHYD 76 76
FT CARBOHYD 127 127
FT CARBOHYD 201 201
FT CARBOHYD 227 227
SQ SEQUENCE 270 AA; 29314 MW; 1FD655A2C4E550B3 CRC64;

Query Match 38.7%; Score 468; DB 1; Length 270;
Best Local Similarity 43.8%; Pred. No. 1.8e-31;
Matches 103; Conservative 47; Mismatches 63; Indels 22; Gaps 8;

QY 1 AOSVSFTFTKFDSDOKDLMEFGHT-ISSNNVIQLKLDNSGNPNVSTSVGRVLYSAPLRW 59
DB 16 AEILSFSFPKFSVQEDLLQGDALVSSGEGLQLTVE--NGVPVWNSTGRALYAPVHIW 74
QY 60 ESST-VVSTFTFTTFOISTPTPTSPGDLAPFLAPYDVTVPNSAGNLLGLFPNLNLR 118
DB 75 DNSTGRVASFATSEFVVKAPVASKSADGIAFLAPLNQI-HGAGGGLYGLF----- 126
QY 119 NSTSKETITIDVNAASNNVAVEFTYDNDIGDPYRKHIGIDVNSIRKATVANDWQNG 178
DB 127 NSSS-----YSSQYQIVAVEFTDTH--TNADPNTRHIGIDVNSIRKATVANDWQNG 176
QY 179 KATAHISYNSAKRSLVTFYPGKK-AVSLSHDVLTQVLQVLPQWIRVGFSASTGL 232
DB 177 EVANVLLTYQAATEMLVSLYTPSNQTSYLSAADVLSKILPEWVRVGFATATGL 231

RESULT 11
LECS_SOPJA STANDARD; PRT; 292 AA.
AC P93535;
AT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Seed lectin precursor (LECSJASG).
OS Sophora japonica (Japanese pagoda tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Sophora.
OX NCBI_TaxID=3897;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=97201486; PubMed=904972;
RA van Damme E.J., Barre A., Rouge P., Peumans W.J.;
RT "Molecular cloning of the bark and seed lectins from the Japanese
RL pagoda tree (Sophora japonica).";
RL Plant Mol. Biol. 33:523-536(1997).
CC -!- FUNCTION: MANNANOSE/GLUCOSE-SPECIFIC LECTIN.
CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or *send an email to license@isb-sib.ch).
CC -----
CC EMBL; U63011; AAB51441.1; -.

```

```

DR HSSP; P19588; ILUL.
DR GlycoSuiteDB; P93535; -.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
KW LECTIN; Calcium; Manganese; Glycoprotein; Signal.
FT SIGNAL 1 37
FT CHAIN 38 292
FT METAL 163 163
FT METAL 165 165
FT METAL 167 167
FT METAL 169 169
FT METAL 172 172
FT METAL 177 177
FT CARBOHYD 82 82
FT CARBOHYD 154 154
FT CARBOHYD 186 186
SQ SEQUENCE 292 AA; 31656 MW; A7431C29117A503E CRC64;

Query Match 37.6%; Score 455; DB 1; Length 292;
Best Local Similarity 43.3%; Pred. No. 2.3e-30;
Matches 104; Conservative 41; Mismatches 63; Indels 32; Gaps 9;

QY 1 AOSVSFTFTKFDSDOKDLMEFGHT-ISSNNVIQLKLDNSGNPNVSTSVGRVLYSAPLRW 59
DB 38 AEILSFSFPKFSVQEDLLQGDALVSSGEGLQLTVE--NGVPVWNSTGRALYAPVHIW 96
QY 60 ESST-VVSTFTFTTFOISTPTPTSPGDLAPFLAPYDVTVPNSAGNLLGLFPNLNLR 113
DB 97 DKSTGRVASFATSEFVVKAPVASKSADGIAFLA-----PPNNIQGPGGHLGLFHS 150
QY 114 LNLRNSTSKETITIDVNAASNNVAVEFTYDNDIGDPYRKHIGIDVNSIRKATVAM 173
DB 151 SG--YNS-----SYQIIAVDFDTH--INAWDPNTRHIGIDVNSIRKATVAM 193
QY 174 DWONGKTATAHISYNSAKRSLVTFYPGKK-AVSLSHDVLTQVLQVLPQWIRVGFSASTGL 232
DB 194 GWONGEVANVLLSYQAATEMLVSLYTPSNQTSYLSAADVLSKILPEWVRVGFATATGL 253

RESULT 12
PHAM_PHAVU STANDARD; PRT; 273 AA.
AC P15231;
AT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Leucoagglutinating phytohemagglutinin precursor (PHA-L).
GN PLEC2.
OS Phaseolus vulgaris (Kidney bean) (French bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
OX NCBI_TaxID=3885;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. PINTO U1111; TISSUE=Leaf;
RA Voelker T.A., Staswick P., Chrispeels M.J.;
RT "Molecular analysis of two phytohemagglutinin genes and their
RL expression in Phaseolus vulgaris cv. Pinto, a lectin-deficient
RL cultivar of the bean.";
RL EMBO J. 5:3075-3082(1986).
CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC COMBINATION IS THE ONLY ONE DEVOID OF AGGLUTINATION ACTIVITY.  
 CC ISOFORM B4 DISPLAYS MAXIMAL AGGLUTINATION ACTIVITY.  
 CC -!- TISSUE SPECIFICITY: STRONG EXPRESSION IN SEED. LOWER LEVELS IN THE  
 CC FLOWER, AND THE BARK OF THE ROOTS. NO EXPRESSION IN LEAF. THE  
 CC LECTIN ACCUMULATES IN THE INNER BARK IN AUTUMN.  
 CC -!- SIMILARITY: TO OTHER LEGUMINOUS LECTINS.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U12782; AAA80181.1; -  
 CC HSSP; P19588; ILUL.  
 CC InterPro; IPR000985; Lectin\_legA.  
 CC InterPro; IPR001220; Lectin\_legB.  
 CC Pfam; PF00138; lectin\_legA; 1.  
 CC Pfam; PF00139; lectin\_legB; 1.  
 CC ProDom; PD000671; Lectin\_legA; 1.  
 CC ProDom; PD000711; Lectin\_legB; 1.  
 CC PROSITE; PS00307; LECTIN\_LEGUME\_BETA; 1.  
 CC PROSITE; PS00308; LECTIN\_LEGUME\_ALPHA; 1.  
 CC LECTIN; Calcium; Manganese; Glycoprotein; Signal; Multigene family.  
 KW SIGNAL 1 31  
 FT CHAIN 1 32 285 BARK AGGLUTININ I, POLYPEPTIDE A.  
 FT METAL 156 158 MANGANESE (BY SIMILARITY).  
 FT METAL 158 158 MANGANESE AND CALCIUM (BY SIMILARITY).  
 FT METAL 160 160 CALCIUM (BY SIMILARITY).  
 FT METAL 162 162 CALCIUM (BY SIMILARITY).  
 FT METAL 166 166 MANGANESE AND CALCIUM (BY SIMILARITY).  
 FT METAL 171 171 MANGANESE (BY SIMILARITY).  
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 285 AA; 30928 MW; 49382E50EEF27282 CRC64;

Query Match 36.9%; Score 446; DB 1; Length 285;  
 Best Local Similarity 42.6%; Pred. No. 1.2e-29;  
 Matches 100; Conservative 47; Mismatches 66; Indels 22; Gaps 8;

QY 3 SVSFTFTKFDSDQKDLMFQGH-TISSNVITQLTKLDSNGNPVSTVGRLVYSAPLRLWES 61  
 DB 34 SLSFEPKFPAPQPYLIFQDALVTSTGVQLQTNV-VNGVPSGKSLGRALYAAPFQIWD 92  
 QY 62 ST-VVSTFTFTTFTTQISTPTTSPGDLAFPLAPDYVIPNSAGNLLGLFPNLNLRNS 120  
 DB 93 TTGNVASFVTSFIIQAPNPTTADGLAFPLAPVD--QPLDVGGMGLGIF-----KDG 144  
 QY 121 TTSKETTIDVNAASNNVAVFEDTYPNDNIGDPYRKHGIDVNSIRKATVANDWQNGKT 180  
 DB 145 YFNK-----SNOIVAVEFDTFSNIHF-DPKGRHNGINVSIVSIKIVPNWNTGEV 194  
 QY 181 ATAHISYNASKRLSVTTTFYPPGKAVSLSHD-VELTQVLQPWIRVGSASTGLEK 234  
 DB 195 ANVFISYEASTKSLASLVYPSLTSFIVHIVDKVDLPENWVRFGFSATTGIDK 249

RESULT 14  
 LCBL\_ROBPS  
 ID LCBL\_ROBPS STANDARD; PRT; 285 AA.  
 AC Q41159;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Seed agglutinin I precursor (RPSAI) (LECRPASI).  
 OS Robinia pseudoacacia (Black locust).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Robinieae; Robinia.  
 CC NCBI\_TaxID=35938;  
 RN [1]

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC COMBINATION IS THE ONLY ONE DEVOID OF AGGLUTINATION ACTIVITY.  
 CC ISOFORM B4 DISPLAYS MAXIMAL AGGLUTINATION ACTIVITY.  
 CC -!- TISSUE SPECIFICITY: STRONG EXPRESSION IN SEED. LOWER LEVELS IN THE  
 CC FLOWER, AND THE BARK OF THE ROOTS. NO EXPRESSION IN LEAF. THE  
 CC LECTIN ACCUMULATES IN THE INNER BARK IN AUTUMN.  
 CC -!- SIMILARITY: TO OTHER LEGUMINOUS LECTINS.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U12782; AAA80181.1; -  
 CC HSSP; P19588; ILUL.  
 CC InterPro; IPR000985; Lectin\_legA.  
 CC InterPro; IPR001220; Lectin\_legB.  
 CC Pfam; PF00138; lectin\_legA; 1.  
 CC Pfam; PF00139; lectin\_legB; 1.  
 CC ProDom; PD000671; Lectin\_legA; 1.  
 CC ProDom; PD000711; Lectin\_legB; 1.  
 CC PROSITE; PS00307; LECTIN\_LEGUME\_BETA; 1.  
 CC PROSITE; PS00308; LECTIN\_LEGUME\_ALPHA; 1.  
 CC LECTIN; Calcium; Manganese; Glycoprotein; Signal; Multigene family.  
 KW SIGNAL 1 31  
 FT CHAIN 1 32 285 BARK AGGLUTININ I, POLYPEPTIDE A.  
 FT METAL 156 158 MANGANESE (BY SIMILARITY).  
 FT METAL 158 158 MANGANESE AND CALCIUM (BY SIMILARITY).  
 FT METAL 160 160 CALCIUM (BY SIMILARITY).  
 FT METAL 162 162 CALCIUM (BY SIMILARITY).  
 FT METAL 166 166 MANGANESE AND CALCIUM (BY SIMILARITY).  
 FT METAL 171 171 MANGANESE (BY SIMILARITY).  
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 285 AA; 30928 MW; 49382E50EEF27282 CRC64;

Query Match 36.9%; Score 446; DB 1; Length 285;  
 Best Local Similarity 42.6%; Pred. No. 1.2e-29;  
 Matches 100; Conservative 47; Mismatches 66; Indels 22; Gaps 8;

QY 3 SVSFTFTKFDSDQKDLMFQGH-TISSNVITQLTKLDSNGNPVSTVGRLVYSAPLRLWES 61  
 DB 34 SLSFEPKFPAPQPYLIFQDALVTSTGVQLQTNV-VNGVPSGKSLGRALYAAPFQIWD 92  
 QY 62 ST-VVSTFTFTTFTTQISTPTTSPGDLAFPLAPDYVIPNSAGNLLGLFPNLNLRNS 120  
 DB 93 TTGNVASFVTSFIIQAPNPTTADGLAFPLAPVD--QPLDVGGMGLGIF-----KDG 144  
 QY 121 TTSKETTIDVNAASNNVAVFEDTYPNDNIGDPYRKHGIDVNSIRKATVANDWQNGKT 180  
 DB 145 YFNK-----SNOIVAVEFDTFSNIHF-DPKGRHNGINVSIVSIKIVPNWNTGEV 194  
 QY 181 ATAHISYNASKRLSVTTTFYPPGKAVSLSHD-VELTQVLQPWIRVGSASTGLEK 234  
 DB 195 ANVFISYEASTKSLASLVYPSLTSFIVHIVDKVDLPENWVRFGFSATTGIDK 249

RESULT 14  
 LCBL\_ROBPS  
 ID LCBL\_ROBPS STANDARD; PRT; 285 AA.  
 AC Q41159;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Seed agglutinin I precursor (RPSAI) (LECRPASI).  
 OS Robinia pseudoacacia (Black locust).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Robinieae; Robinia.  
 CC NCBI\_TaxID=35938;  
 RN [1]

```

RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-51.
RX TISSUE-Seed;
RC MEDLINE-96191285; PubMed-8616218;
RA van Damme E.J.M., Barre A., Rouge P., van Leuven F., Peumans W.J.;
RT "The seed lectins of black locust (Robinia pseudoacacia) are encoded
RT by two genes which differ from the bark lectin genes.";
RL Plant Mol. Biol. 29:1197-1210(1995).
CC -!- FUNCTION: SEED LECTIN.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN SEED.
CC -!- PTM: THREE OF THE FOUR POTENTIAL GLYCOSYLATION SITES ARE OCCUPIED.
CC -!- SIMILARITY: TO OTHER LEGUMINOUS LECTINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U24250; AAC49272.1;
CC HSSP; P19588; ILUL.
CC InterPro: IPR000985; Lectin_legA.
CC InterPro: IPR001220; Lectin_legB.
CC Pfam; PF00138; lectin_legA; 1.
CC Pfam; PF00139; lectin_legB; 1.
CC ProDom; PD000671; Lectin_legA; 1.
CC ProDom; PD000711; Lectin_legB; 1.
CC PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
CC PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
CC LECTIN; Calcium; Manganese; Glycoprotein; Signal; Multigene family.
CC SIGNAL 1 31
CC CHAIN 32 285 SEED AGGLUTININ I.
CC METAL 156 156 MANGANESE AND CALCIUM (BY SIMILARITY).
CC METAL 158 158 MANGANESE AND CALCIUM (BY SIMILARITY).
CC METAL 162 162 CALCIUM (BY SIMILARITY).
CC METAL 166 166 MANGANESE AND CALCIUM (BY SIMILARITY).
CC METAL 171 171 MANGANESE (BY SIMILARITY).
CC CARBOHYD 92 92 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 147 147 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 188 188 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 209 209 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 285 AA; 30943 MW; 6AE82CDC920224CE CRC64;

Query Match 35.8%; Score 445; DB 1; Length 285;
Best Local Similarity 41.9%; Pred. No. 1.5e-29;
Matches 98; Conservative 49; Mismatches 65; Indels 22; Gaps 8;

QY 3 SVSFTFTKDSQDKLDMFOGHT-ISSNNVLTQLKDSNGNPVSTSVGRVLYSAPRLWES 61
DB 34 SLSEFPKFAFPNQYLIFQRDALVTSTGVQLTNN-VNGVPPRRSIGRSLYAAPFQIWDN 92

QY 62 ST-VVSTFTFTTFOISTPYTSPPGDGLAFFLAPVDYVIPPNSAGNLLGLFPNLNLRNS 120
DB 93 TTGNVASFVTSFSLIQAPNPATTADGLAFFLAPVDY-QPGDLGGMLGIF-----KDG 144

QY 121 TTSKETIDVNAASNNVAVEDFYDNPNDIGDPYRKHIGIDVNSRKSATVAMWQNGKT 180
DB 145 SYNK-----SNQIVAVEFDYFSNIHF-DPKGRHMGINNVNSIVSKTVPNWNTNGEV 194

QY 181 ATAHISYNSAKRLSVTTFYPGGKAVLSHD-VELTOVLPOWIRVGFSASTGLE 233
DB 195 ANVFISYEASTKSLNASLYPSLSTFSFIHAIVDVKVLPWVRFVGFSAATTGID 248

RESULT 15
LCS2_ROBPS STANDARD; PRT; 285 AA.
AC Q41161.
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)

```

```

DE SEQUENCE FROM N.A., AND SEQUENCE OF 32-50.
OS Robinia pseudoacacia (Black locust).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Erosids 1; Fabales; Fabaceae; Papilionoideae; Robinieae; Robinia.
OX NCBI_TaxID=35938;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-50.
RC TISSUE-Seed;
RX MEDLINE-96191285; PubMed-8616218;
RA van Damme E.J.M., Barre A., Rouge P., van Leuven F., Peumans W.J.;
RT "The seed lectins of black locust (Robinia pseudoacacia) are encoded
RT by two genes which differ from the bark lectin genes.";
RL Plant Mol. Biol. 29:1197-1210(1995).
CC -!- FUNCTION: SEED LECTIN.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- TISSUE SPECIFICITY: EXPRESSION IN SEED.
CC -!- PTM: MOSTLY FOUND IN NON-GLYCOSYLATED FORM.
CC -!- SIMILARITY: TO OTHER LEGUMINOUS LECTINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U24249; AAC49271.1;
CC HSSP; P19588; ILUL.
CC InterPro: IPR000985; Lectin_legA.
CC InterPro: IPR001220; Lectin_legB.
CC Pfam; PF00138; lectin_legA; 1.
CC Pfam; PF00139; lectin_legB; 1.
CC ProDom; PD000671; Lectin_legA; 1.
CC ProDom; PD000711; Lectin_legB; 1.
CC PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
CC PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
CC LECTIN; Calcium; Manganese; Glycoprotein; Signal; Multigene family.
CC SIGNAL 1 31
CC CHAIN 32 285 SEED AGGLUTININ II.
CC METAL 156 156 MANGANESE (BY SIMILARITY).
CC METAL 158 158 MANGANESE AND CALCIUM (BY SIMILARITY).
CC METAL 162 162 CALCIUM (BY SIMILARITY).
CC METAL 166 166 MANGANESE AND CALCIUM (BY SIMILARITY).
CC METAL 171 171 MANGANESE (BY SIMILARITY).
CC CARBOHYD 147 147 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 285 AA; 31021 MW; 2C0B3249620294DE CRC64;

Query Match 36.7%; Score 444; DB 1; Length 285;
Best Local Similarity 42.7%; Pred. No. 1.8e-29;
Matches 100; Conservative 46; Mismatches 66; Indels 22; Gaps 8;

QY 3 SVSFTFTKDSQDKLDMFOGHT-ISSNNVLTQLKDSNGNPVSTSVGRVLYSAPRLWES 61
DB 34 SLSEFPKFAFPNQYLIFQRDALVTSTGVQLTNN-VNGVPPRRSIGRSLYAAPFQIWDN 92

QY 62 ST-VVSTFTFTTFOISTPYTSPPGDGLAFFLAPVDYVIPPNSAGNLLGLFPNLNLRNS 120
DB 93 TTGNVASFVTSFSLIQAPNPATTADGLAFFLAPVDY-QPLDLGGMLGIF-----KNG 144

QY 121 TTSKETIDVNAASNNVAVEDFYDNPNDIGDPYRKHIGIDVNSRKSATVAMWQNGKT 180
DB 145 YFNK-----SNQIVAVEFDYFSNRH-WDPTGRHMGINNVNSIVSKTVPNWNTNGEV 194

QY 181 ATAHISYNSAKRLSVTTFYPGGKAVLSHD-VELTOVLPOWIRVGFSASTGLE 233
DB 195 ANVFISYEASTKSLNASLYPSLSTFSFIHAIVDVKVLPWVRFVGFSAATTGID 248

Search completed: February 26, 2003, 16:52:56
Job time : 6.7191 secs

```

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:43:23 ; Search time 22.2022 Seconds  
(without alignments)  
2171.628 Million cell updates/sec

Title: US-09-476-485A-8  
Perfect score: 1209  
Sequence: 1 AQSVSFTFKDSQKDLMF.....TQVLQWIRVGFSAVGLPK 234

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	796.5	65.9	272	10 Q9ZTA9	Q9ZTA9 dolichos la
2	776.5	64.2	279	10 Q9M7M4	Q9M7M4 phaseolus v
3	649.5	53.7	290	10 Q04672	Q04672 canavalia b
4	643.5	53.2	290	10 Q947H0	Q947H0 canavalia e
5	547	45.2	266	10 P93536	P93536 sophora jap
6	544	45.0	284	10 Q9FVU9	Q9FVU9 sophora fla
7	532	44.0	293	10 P93537	P93537 sophora fla
8	488.5	40.4	254	10 Q43376	Q43376 arachis hyp
9	485.5	40.2	280	10 Q43374	Q43374 arachis hyp
10	480.5	39.7	282	10 P93247	P93247 maackia amu
11	480.5	39.7	286	10 P93248	P93248 maackia amu
12	477.5	39.5	254	10 Q43377	Q43377 arachis hyp
13	459	38.0	258	10 Q9FVF8	Q9FVF8 ulex europe
14	457	37.8	285	10 Q9ZNP6	Q9ZNP6 robinia pse
15	449.5	37.2	279	10 Q49899	Q49899 medicago sa
16	449	37.1	256	10 P93246	P93246 maackia amu

17	446	36.9	275	10 Q8RVY1	Q8RVY1 phaseolus v
18	444.5	36.8	273	10 Q8RVY4	Q8RVY4 phaseolus c
19	443.5	36.7	275	10 Q8RVH1	Q8RVH1 phaseolus v
20	438.5	36.3	276	10 Q40750	Q40750 phaseolus a
21	436	36.1	274	10 Q43628	Q43628 phaseolus v
22	434.5	35.9	275	10 Q8RVX5	Q8RVX5 phaseolus v
23	433.5	35.9	273	10 Q42411	Q42411 medicago sa
24	433.5	35.9	273	10 Q8RVH2	Q8RVH2 phaseolus v
25	427.5	35.4	275	10 Q8RW23	Q8RW23 phaseolus c
26	427.5	35.4	275	10 Q8RVX6	Q8RVX6 phaseolus v
27	427.5	35.4	275	10 Q8RVH3	Q8RVH3 phaseolus v
28	426.5	35.3	275	10 Q8RVY0	Q8RVY0 phaseolus v
29	424.5	35.1	278	10 Q9LED9	Q9LED9 phaseolus l
30	424.5	35.1	278	10 Q9LED8	Q9LED8 phaseolus l
31	421.5	34.9	251	10 Q8WIR7	Q8WIR7 griffonia s
32	421	34.8	285	10 Q9ZWP5	Q9ZWP5 robinia pse
33	418.5	34.6	247	10 Q8WIR6	Q8WIR6 griffonia s
34	417.5	34.5	251	10 Q93X48	Q93X48 lens ervoid
35	417.5	34.5	251	10 Q93WH6	Q93WH6 lens culina
36	417.5	34.5	251	10 Q8RVX9	Q8RVX9 phaseolus v
37	416.5	34.4	251	10 Q93X49	Q93X49 lens culina
38	415.5	34.4	251	10 Q93X41	Q93X41 lens culina
39	415.5	34.4	251	10 Q8W4Y4	Q8W4Y4 lens lamott
40	415.5	34.4	255	10 Q40912	Q40912 phaseolus l
41	414.5	34.3	251	10 Q8VXF2	Q8VXF2 lens culina
42	414.5	34.3	268	10 P93458	P93458 phaseolus l
43	414	34.2	240	10 Q9SM56	Q9SM56 psophocarpu
44	413.5	34.2	251	10 Q93X50	Q93X50 lens culina
45	412.5	34.1	251	10 Q8RW33	Q8RW33 lathyrus sa

## ALIGNMENTS

### RESULT 1

Q9ZTA9 PRELIMINARY; PRT; 272 AA.  
ID Q9ZTA9  
AC Q9ZTA9;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Mannose lectin.  
GN FRIL.  
OS Dolichos lab lab (field bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Dolichos.  
OX NCBI\_TaxID=35936;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=COTYLEDON;  
RX MEDLINE=99110944; PubMed=9892687;  
RA Colucci G., Moore J.G., Feldman M., Chrispeels M.J.;  
RT "CDNA cloning of FRIL, a lectin from Dolichos lablab, that preserves  
RT hematoepoietic progenitors in suspension culture.";  
RT Proc. Natl. Acad. Sci. U.S.A. 96:646-650(1999).  
DR EMBL; AF067417; AAD10734.1;  
DR HSSP; P02866; IONA.  
DR InterPro; IPR000985; Lectin\_legA.  
DR InterPro; IPR001220; Lectin\_legB.  
DR Pfam; PF00138; Lectin\_legA\_1.  
DR Pfam; PF00139; Lectin\_legB\_1.  
DR ProDom; PD000671; Lectin\_legA; 1.  
DR ProDom; PD000711; Lectin\_legB; 1.  
DR PROSITE; PS00308; LECTIN\_LEGUME\_ALPHA; 1.  
DR PROSITE; PS00307; LECTIN\_LEGUME\_BETA; UNKNOWN\_1.  
KW Lectin.  
SQ SEQUENCE 272 AA; 29900 MW; EA6C004307441495 CRC64;

Query Match 65.9%; Score 796.5; DB 10; Length 272;  
Best Local Similarity 66.7%; Pred. No. 6.7e-57;  
Matches 156; Conservative 30; Mismatches 45; Indels 3; Gaps 2;

```

QY 1 AQSVSFTFTKFDSDOKLDMFOGHITSSNNVQLTKLDSNGPNVSTSVGRVLYSAPLRLWE 60
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 9 AQSLSFSTFKFDPNQEDLIIFQGHATSNVNLQVTKLDSAGNPNVSSAGRVLYSAPLRLWE 68
QY 61 SSTVVSFTFTFTQISTPTSPPGDGLAFPLADYDVIPPNNSAGNLLGLFPNLNLRNS 120
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 69 DSAVLTSFTDIINFEISTPTSTRADGLAFFIAPDPSVI--SYHGGLGLFPNANTLNNS 126
QY 121 TTSKETTIDVNAASN--NVVAVERTYDNDTGDYKHKHIGIDVNSIRSKATVAVDWQNGKT 180
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 127 STSENQT-TTKAASSNVVAVFEDYLYLPDYGDPNKHIGIDVNSIRSKATVAKDWQNGKI 185
QY 181 ATAHISYNSAKRLSVTTFYPPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 186 ATAHISYNSKRLSVTSYYAGSKPATLSYDIELHTLVPWVRVGLSASTGQDK 239

RESULT 2
Q9M7M4 PRELIMINARY; PRT; 279 AA.
AC Q9M7M4;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
DE Mannose lectin FRIL (Fragment).
OS Phaseolus vulgaris (Kidney bean) (French bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
OX NCBI_TaxID=3885;
RN [1]
RP SEQUENCE FROM N.A.
RA Moore J.G., Colucci G., Fuchs C.A., Hicklin D.J., Chrispeels M.J.,
RA Feldman M.;
RT "A new lectin in red kidney bean called PvFRIL stimulates
RT proliferation of NIH3T3 cells expressing the Flt3 receptor.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF121458; AAF28739.1; -.
DR HSSP; P02866; IONA.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; lectin_legA; 1.
DR ProDom; PD000711; lectin_legB; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
KW Lectin.
FT NON_TER 1
SQ SEQUENCE 279 AA; 31102 MW; F8919CF8B3EE4652 CRC64;

Query Match 64.2%; Score 776.5; DB 10; Length 279;
Best Local Similarity 66.8%; Pred. No. 2.9e-55;
Matches 157; Conservative 25; Mismatches 48; Indels 5; Gaps 4;

QY 1 AQSVSFTFTKFDSDOKLDMFOGHITSSNNVQLTKLDSNGPNVSTSVGRVLYSAPLRLWE 60
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 AQSLSFNTKFDLDOKLIFQGDATSTNNVQLTKLDSGNGPNVGSVGRVLYSAPFLWE 60
QY 61 SSTVVSFTFTFTQISTPTSPPGDGLAFPLADYDVIPPNNSAGNLLGLFPNLNLRNS 120
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 NSMAVSFETNLTIQISTPHPYAAGDGAFFLADHTVIPPNSWKGFLGYSNV--FRNS 118
QY 121 TTSKETTIDVNAASN--NVVAVERTYDNDTGDYKHKHIGIDVNSIRSKATVAVDWQNGK 179
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 119 PTSENQSGDVNTDS--RVVAVERTYDNPANI--DPNRYRHIGIDVNSIKSETARWEQNGK 176
QY 180 TATAHISYNSAKRLSVTTFYPPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 177 TATARISYNSAKSKSTVTFYPPGMEVVALSHDVELTQVLPQWIRVGFSASTGLEK 231

RESULT 3

```

```

004672 PRELIMINARY; PRT; 290 AA.
AC Q04672;
DT 01-JUL-1997 (T-EMBLrel. 04, Created)
DT 01-JUL-1997 (T-EMBLrel. 04, Last sequence update)
DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
DE Lactin (Fragment).
OS Canavalia brasiliensis (Brazilian jack bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Canavalia.
OX NCBI_TaxID=61861;
RN [1]
RP SEQUENCE FROM N.A.
RA Grangeiro T.B.;
RL Thesis (1996), Universidade Federal do Ceara, Fortaleza, Brazil.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97454283; PubMed=9310358;
RA Grangeiro T.B., Schriefer A., Calvete J.J., Raida M., Urbanke C.,
RA Barral-Netto M., Cavada B.S.;
RT "Molecular cloning and characterization of ConBr, the lectin of
RT Canavalia brasiliensis seeds.";
RL Eur. J. Biochem. 248:43-48(1997).
DR EMBL; Y13904; CAA74202.1; -.
DR HSSP; P02866; IONA.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; lectin_legA; 1.
DR ProDom; PD000711; lectin_legB; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
FT NON_TER 290
SQ SEQUENCE 290 AA; 31408 MW; 0CE5B4BBC950AC92 CRC64;

Query Match 53.7%; Score 649.5; DB 10; Length 290;
Best Local Similarity 54.5%; Pred. No. 6.2e-45;
Matches 128; Conservative 35; Mismatches 63; Indels 9; Gaps 4;

QY 3 SVSFTFTKFDSDOKLDMFOGH--TISSNNVQLTKLDSNGPNVSTSVGRVLYSAPLRLWE 61
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 36 ALHFENQFSKQKDLILQGDATGTGDNLELTRVSSNGSPQGSVGRALFYAPVHIWES 95
QY 62 STVVSFTFTFTQISTPTSPPGDGLAFPLADYDVIPPNNSAGNLLGLFPNLNLRNST 121
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 96 SAVVASFEATFTELIKSP--DSHPADGIAFTISNIDSSIPSGTGLLGLFPDANVRNS- 153
QY 122 TSKETTIDVNAASN--NVVAVERTYDNDTGDYKHKHIGIDVNSIRSKATVAVDWQNGK 179
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 154 ----TTIDFNAAYNADTIVAVELDTYPTNDIGDPSYPHIGIDIKSVRSKKTAKNMQNGK 209
QY 180 TATAHISYNSAKRLSVTTFYPPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 210 VCTAHIIYNSVGRKLSAVSVYPNGDSATVSDVDLDNVLPWVRVGLSASTGLYK 264

RESULT 4
Q947H0 PRELIMINARY; PRT; 290 AA.
AC Q947H0;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
DE Concanavalin A.
GN CONA.
OS Canavalia ensiformis (Jack bean) (Horse bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Canavalia.
OX NCBI_TaxID=3823;
RN [1]

```



DR PROSITE; PS00308; LECTIN\_LEGUME\_ALPHA; 1.  
DR PROSITE; PS00307; LECTIN\_LEGUME\_BETA; UNKNOWN\_1.  
KW Lectin; Glycoprotein; Signal; Mannose-binding; Multigene family.  
FT SIGNAL 1 35  
FT CHAIN 36 162 BARK LECTIN I, A1 SUBUNIT.  
FT CHAIN 163 293 BARK LECTIN I, B1 SUBUNIT.  
FT CARBOHYD 64 64 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 152 152 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 213 213 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CONFLICT 57 R -> S (IN REF. 2).  
SQ SEQUENCE 293 AA; 33221 MW; 3D2F191AD63F1986 CRC64;  
Query Match 44.0%; Score 532; DB 10; Length 293;  
Best Local Similarity 47.5%; Pred. No. 2.le-35;  
Matches 112; Conservative 42; Mismatches 66; Indels 16; Gaps

---

QY 1 AQSVFTTFFKDSQDKLMPQGHTISSNNVIQLTKLDNSNGNPYSTSGRVLYSAPRLWE 60  
DB :|||:::||::|||::|||::|||::|||::|||::|||::|||::|||:  
36 SDLSFTVENFPNPDILLORDASITSNELLTLRTSNGKPQKSVGRAIYPVRLWD 95  
QY 61 SST-VVSFTPTTFEQISTSYTPSPGDGLAFFLPADPYTVPPNSAGNLGLGFPLNALRN 119  
DB ||::::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
96 KSTGRLASFTSFYSVTSP-TTDPGDIATFIAPPDTT--PGYTGGLLGLF-----N 145  
QY 120 STTSKETTTIDVAASN---VVAVEETTPDNIGIPYRKHKIGIDVNSIRSKATAVDW 175  
DB |:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
146 SSTVSSNSSDHGVAFHNSLPQIVAVEEDTYING-RDPNRHHVIDVNSIKRSTTKWTW 204  
QY 176 QNGKTATAHIYSANSKRSLSTVFYFGKAVSLSHDWELTQVLFWIRVGSASTG 231  
DB ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
205 RNGVEATANISYNPVSQLTAVSSPNSEPIITHYDIDLKTLVPKWRVYGFSASTG 260  
LEC.

### RESULT 8

ID	QA33376	PRELIMINARY;	PRT;	254 AA.
DC	AQ43376;			
AD	01-NOV-1996 (TrEMBLrel. 01, Created)			
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)			
DE	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)			
DE	Mannose/glucose-binding lectin precursor (Fragment).			
OS	Arachis hypogaea (peanut).			
GC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;			
OC	eurosidia I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;			
OX	Arachis.			
OX	NCBI_TaxID=3818;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. SELLIE; TISSUE=SEED;			
RA	Law J.J.;			
RT	"Cloning and expression of cDNA for mannose-binding lectin from			
RT	peanut."			
RL	Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.			
DR	EMBL; U22472; AAA74575.1; ..			
DR	HSSP; P02867; 2BQP.			
DR	InterPro; IPR000985; Lectin_legA.			
DR	InterPro; IPR001220; Lectin_legB.			
DR	Pfam; PF00138; lectin_legA; 1.			
DR	Pfam; PF00139; lectin_legB; 1.			
DR	ProDom; PD000671; Lectin_legA; 1.			
DR	ProDom; PD000711; Lectin_legB; 1.			
DR	PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.			
KW	Lectin; Signal.			
FT	NON_TER 1 1			
FT	SIGNAL <1 2			
SQ	SEQUENCE 254 AA; 28202 MW; 3AE1EALIF90BICA03 CRC64;			

Query Match 40.4%; Score 488.5; DB 10; Length 254;  
Best Local Similarity 46.6%; Pred. No. 6e-32;  
Matches 109; Conservative 42; Mismatches 64; Indels 19; Gaps



```

QY 3 SVSFTTKFD-SDOKDLMOFGH-TISSNNVLTQLKDSNGNPVSTSVGRVLYSAPRLWE 60
Db 3 SLFSFNKFEQDDERLILGDDATFSASKGTLTKVDANGTPAKSTVGRVLYHSTQVRLWE 62
QY 61 SST-VVSTFTTFTTQISTPTTSPPGDGLAFFLAPYDTVIPPNSAGNLLGLFPNLNLRN 119
Db 63 KSTNRLTNFQAQSFVLIKSP-NDIGADGIAFFIAAPDSQIPKNSAGGTGLGLFDPTQA--- 118
QY 120 STTSKETTIDVNAASNNVAVFEDT-YPDNIG-DPYRKHGIDVNSTRSKATVANDWQW 177
Db 119 -----QNPANQVLAVEEDTFTAQDSNGWDPNYQHIGIDVNSIKSAATTKWERN 168
QY 178 GKATAHISYNSAKRLSVTFYPPGKAVSLSHDVELTQVLPQWIRVGFSASTG 231
Db 169 GQTLNLVLTIDANSKNLQVTASYPDGQRYQVSYVVDLRLDHPWGRVGFSAAG 222

RESULT 9
ID Q43374 PRELIMINARY; PRT; 280 AA.
AC Q43374;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Mannose/glucose-binding lectin precursor.
GN LEC.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. SELLIE; TISSUE=NODEULE;
RA Law I.J.;
RT "Cloning and expression of cDNA for mannose-binding lectin from
RL peanut".
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U22469; AAA74572.1; -.
DR HSSP; P02866; IDQ2.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; lectin_legA; 1.
DR ProDom; PD000711; lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGME_BETA; 1.
DR PROSITE; PS00307; LECTIN_LEGME_BETA; UNKNOWN_1.
KW Lectin; Signal.
FT CHAIN 1 26 MANNOSE/GLUCOSE-BINDING LECTIN.
FT SIGNAL 27 280
SQ SEQUENCE 280 AA; 31012 MW; C15B39B32F455BD5 CRC64;

Query Match 40.2%; Score 485.5; DB 10; Length 280;
Best Local Similarity 45.7%; Pred. No. 1.2e-31;
Matches 107; Conservative 47; Mismatches 61; Indels 19; Gaps 8;

QY 3 SVSFTTKFD-SDOKDLMOFGH-TISSNNVLTQLKDSNGNPVSTSVGRVLYSAPRLWE 60
Db 29 SLFSFNKFEQDDERLILGDDATFSASKGTLTKVDANGTPAKSTVGRVLYHSTQVRLWE 88
QY 61 SST-VVSTFTTFTTQISTPTTSPPGDGLAFFLAPYDTVIPPNSAGNLLGLFPNLNLRN 119
Db 89 KSTNRLTNFQAQSFVLIKSP-IDN-GADGIAFFIAAPDSQIPKNSAGGTGLGLF-----D 140
QY 120 STTSKETTIDVNAASNNVAVFEDT-YPDNIG-DPYRKHGIDVNSTRSKATVANDWQW 177
Db 141 PSTAQ-----NPSANQVLAVEEDTFTAQDSNGWDPNYQHIGIDVNSIKSAATTKWERN 194
QY 178 GKATAHISYNSAKRLSVTFYPPGKAVSLSHDVELTQVLPQWIRVGFSASTG 231
Db 195 GQTLNLVLTIDANSKNLQVTASYPDGQRYQVSYVVDLRLDLPWGRVGFSAAG 248

```

```

RESULT 10
P93247 PRELIMINARY; PRT; 282 AA.
AC P93247;
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Bark LEUCOAGGLUTININ I precursor (MALBI) (Fragment).
GN LECMALBI.
OS Maackia amurensis.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Maackia.
OX NCBI_TaxID=37501;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=BARK;
RX MEDLINE=97390228; PubMed=9249142;
RA van Damme E.J.M.; Van Leuven F.; Peumans W.J.;
RT "Isolation, characterization and molecular cloning of the bark lectins
RT from Maackia amurensis.";
RL Glycoconj. J. 14:449-456(1997).
CC -!- FUNCTION: BARK LECTINS ARE STORAGE PROTEIN THAT PROBABLY MAINTAINS
CC STOCKS OF NITROGEN DURING DORMANT PERIOD. SELF-AGGREGATABLE
CC MOLECULES THAT CAN BIND THEIR OWN CARBOHYDRATE SIDE CHAINS. THEY
CC COULD ALSO PLAY A ROLE IN THE PLANT'S DEFENCE AGAINST PHYTOPHAGOUS
CC INVERTEBRATES OR HERBIVOROUS HIGHER ANIMALS (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTETRAMER OR HETEROTETRAMER OF MAHB AND MALB SUBUNITS.
CC -!- PTM: ONLY THREE OF THE PUTATIVE GLYCOSYLATION SITES ARE OCCUPIED.
CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
DR EMBL; U65009; AAB39933.1; -.
DR HSSP; P19588; LLUL.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR InterPro; IPR000531; TonB_boxC.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; lectin_legA; 1.
DR ProDom; PD000711; lectin_legB; 1.
DR PROSITE; PS00308; LECTIN_LEGME_ALPHA; 1.
DR PROSITE; PS00307; LECTIN_LEGME_BETA; UNKNOWN_1.
DR PROSITE; PS00430; TONE_DEPENDENT_REC_1; UNKNOWN_1.
KW Lectin; Calcium; Manganese; Glycoprotein; Multigene family; Signal.
FT NON-TER 1 24 BY SIMILARITY.
FT SIGNAL <1 24
FT CHAIN 25 282 BARK LEUCOAGGLUTININ I.
FT CARBOHYD 23 23 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 282 AA; 45B3F714E1D8957A CRC64;

Query Match 39.7%; Score 480.5; DB 10; Length 282;
Best Local Similarity 45.3%; Pred. No. 3.1e-31;
Matches 107; Conservative 41; Mismatches 67; Indels 21; Gaps 7;

QY 1 AQVSFTTKFDSDOKDLMOFGH-TISSNNVLTQLKDSNGNPVSTSVGRVLYSAPRLW 59
Db 25 SDELSFTINNFLPNEADLLFOGEASVSTGVLTQTRVE-NGQPKQYSVGRALYAAPRVW 83
QY 60 ESST-VVSTFTTFTTQISTPTTSPPGDGLAFFLAPYDTVIPPNSAGNLLGLFPNLNLR 118
Db 84 DNTTGSVASEFTSTFTFYVKAPNPITNGLAFFLAPPDSQIPTGVTYGLGFNN----- 138
QY 119 NSTTSKETTIDVNAASNNVAVFEDT-Y-PNDNIGDPYRKHGIDVNSTRSKATVANDWQ 176
Db 139 ---TSDD-----SSNQIIVAVEEDTFLHKYNPDPNVRHIGIDVNGIDSKTQWDWI 188
QY 177 NGKATAHISYNSAKRLSVTFYPPGK-AVSLSHDVELTQVLPQWIRVGFSASTG 231

```





Db 185 PWSLVNKKANVVIGFNGATNVLSDVEYPLVRHHTLSHVPLKDVVPEWVRIGFSAATG 244

QY 232 LE 233

Db 245 DE 246

Search completed: February 26, 2003, 16:54:27  
Job time : 23.2022 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:40:30 ; Search time 28.6292 Seconds  
(without alignments)  
1089.120 Million cell updates/sec

Title: US-09-476-485A-8

Perfect score: 1209

Sequence: 1 AGSVFTFTKFDSDQKDLMF.....TQVLPOWIRGVFSASTGLEK 234

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1193	98.7	234	AA662901	Amino acid sequenc
2	796.5	65.9	264	AA887973	A lectin derived p
3	796.5	65.9	264	AA662890	Amino acid sequenc
4	796.5	65.9	286	AA662894	Alpha-amylase inhi
5	776.5	64.2	303	AA662898	Amino acid sequenc
6	643.5	53.2	290	AAV58736	Jack bean concanav
7	638.5	52.8	290	AAV58737	Jack bean concanav
8	537.5	44.5	237	AA74765	Legume concanavall
9	457	37.8	285	20 AAY06812	R. pseudoacacia le
10	446	36.9	285	20 AAY06811	R. pseudoacacia le

11	445	36.8	285	20 AAY06814	Amino acid sequenc
12	444	36.7	285	20 AAY06815	Amino acid sequenc
13	411.5	34.0	275	21 AAY58738	Pea lectin. Pisum
14	379	31.3	242	15 AAR45911	Bauhinia purpurea l
15	379	31.3	262	20 AAY27265	B. purpurea native
16	371	30.7	242	15 AAR45912	Bauhinia purpurea l
17	368.5	30.5	239	16 AAR74767	Pea lectin, lile.
18	367	30.4	236	16 AAR64807	PNA lectin subunit
19	361.5	29.9	228	16 AAR74766	Pea lectin, 2lth.
20	341	28.2	274	20 AAY06813	R. pseudoacacia le
21	269	22.2	282	21 AAB25452	Pinus radiata cell
22	254	21.0	632	21 AAB25109	Pinus radiata cell
23	249	20.6	669	23 ABB93142	Herbicidally activ
24	244	20.2	265	10 AAP91967	Sequence of arceli
25	238.5	19.7	681	23 ABB93833	Herbicidally activ
26	236	19.5	652	23 ABB93374	Herbicidally activ
27	230	19.0	272	23 ABB91343	Herbicidally activ
28	223.5	18.5	244	10 AAP93640	Sequence of a Phas
29	223	18.4	677	23 ABB92654	Herbicidally activ
30	219	18.1	675	23 ABB92029	Herbicidally activ
31	212.5	17.6	674	23 ABB93878	Herbicidally activ
32	211	17.5	691	23 ABB93881	Herbicidally activ
33	208.5	17.2	715	23 ABB92649	Herbicidally activ
34	207	17.1	283	23 ABB91344	Herbicidally activ
35	206	17.0	627	23 ABB91918	Herbicidally activ
36	202	16.7	623	23 ABB91919	Herbicidally activ
37	199	16.5	685	23 ABB93319	Herbicidally activ
38	196.5	16.3	666	23 ABB91530	Herbicidally activ
39	196.5	16.3	693	23 ABB92247	Herbicidally activ
40	194.5	16.1	688	23 ABB93318	Herbicidally activ
41	194.5	16.1	711	23 ABB93337	Herbicidally activ
42	192	15.9	656	23 ABB91011	Herbicidally activ
43	188	15.6	661	23 ABB92738	Herbicidally activ
44	185.5	15.3	649	23 ABB93134	Herbicidally activ
45	181.5	15.0	664	23 ABB92106	Herbicidally activ

#### ALIGNMENTS

##### RESULT 1

AA662901  
ID AA662901 standard; Protein; 234 AA.

AC AA662901;

DT 17-SEP-2001 (first entry)

DE Amino acid sequence of a partial FRIL polypeptide.

FRIL; FRIL2/Flt3 tyrosine kinase receptor-interacting lectin;

progenitor cell preservation factor; radiotherapy; chemotherapy;

progenitor cell; hemangioblast; mesenchymal stem cell; cancer;

severe combined immunodeficiency; aplastic anemia; tissue repair.

OS Sphenostylis stenocarpa.

PN WO200149851-A1.

XX 12-JUL-2001.

PD 30-DEC-1999; 99WO-US31307.

PR 30-DEC-1999; 99WO-US31307.

PA (PHYL-) PHYLOGIX LLC.

PI Colucci MG, Chrispeels MJ, Moore JG;

XX WPI; 2001-441882/47.

XX Legume Progenitor cell preservation factors for in vivo or ex vivo  
PT preservation of hematopoietic progenitor cells and as therapeutics for

PT alleviating/reducing progenitor cell-depleting activity of cancer  
 XX therapeutics  
 PS  
 XX Example 22; Page 116-117; 173pp; English.

CC The present sequence is a partial a FRIL (Flk2/Flt3 tyrosine kinase  
 CC receptor-interacting lectin) protein. The specification describes a  
 CC composition of one or more members of FRIL family of progenitor cell  
 CC preservation factors. The composition is useful for alleviating or  
 CC reducing the hematopoietic progenitor cell-depleting activity of a  
 CC therapeutic treatment, including radiotherapeutic and/or  
 CC chemotherapeutic treatments. Administration of FRIL compositions to a  
 CC patient prior to treatment of the patient with a therapeutic treatment  
 CC having a hematopoietic progenitor cell-depleting activity alleviates or  
 CC reduces the hematopoietic progenitor cell-depleting activity of the  
 CC therapeutic treatment in the patient. FRIL family members are useful for  
 CC isolating population of progenitor cells, hemangioblasts, and mesenchymal  
 CC stem cells. The composition is administered to reduce progenitor cell  
 CC depleting effects of chemotherapeutics, so that the patient can receive  
 CC a higher dose of the chemotherapeutic and preferably recover from cancer.  
 CC It is also administered to patients having, or predisposed to developing  
 CC a condition where the patients hematopoietic progenitor cells are  
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
 CC The isolated mesenchymal cells are useful for tissue repair.

XX  
 SQ Sequence 234 AA;

Query Match 98.7%; Score 1193; DB 22; Length 234;  
 Best Local Similarity 99.1%; Pred. No. 4.3e-102;  
 Matches 232; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AQSVSFTTFKFDSDQKLMFQGHGTTSSNVLTQLKDSNGNPVSTSVGRVLYSAPLRLWE 60  
 Db 1 AQSVSFTTFKFDSDQKLMFQGHGTTSSNVLTQLKDSNGNPVSTSVGRVLYSAPLRLWE 60

Qy 61 SSTVSVSTFTTFOISTPYTSPGDLGAFFLAPYDVTVPNSAGNLLGLFPNLNLRNS 120  
 Db 61 SSTVSVSTFTTFOISTPYTSPGDLGAFFLAPYDVTVPNSAGNLLGLFPNLNLRNS 120

Qy 121 TTSKETTIDVNAASNVAVEFDTPNDNIGDPYRKHGIDVNSIRSKATVAVDWMQNGKT 180  
 Db 121 TTSKETTIDVNAASNVAVEFDTPNDNIGDPYRKHGIDVNSIRSKATVAVDWMQNGKT 180

Qy 181 ATAHISYNSASKRLSVTTFYPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234  
 Db 181 ATAHISYNSASKRLSVTTFYPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234

RESULT 2  
 AAW87973  
 ID AAW87973 standard; Protein; 264 AA.  
 XX  
 AC AAW87973;  
 XX  
 DT 13-APR-1999 (first entry)  
 XX  
 DE A lectin derived progenitor cell preservation factor.  
 XX  
 KW Lectin derived progenitor cell preservation factor; progenitor cell;  
 KW haematopoietic cell; cultured cell preservation; anticancer therapy;  
 KW myeloblastic therapy; sickle-cell anaemia; ablative therapy protection;  
 XX FLK2/FLT3 receptor.  
 XX  
 OS Dolichos lab lab.  
 XX  
 PN WO9859038-A1.  
 XX  
 PD 30-DEC-1998.  
 XX  
 PF 23-JUN-1998; 98WO-US13046.  
 XX  
 PR 24-JUN-1997; 97US-0881189.  
 XX

PA (IMCL-) IMCLONE SYSTEMS INC.  
 XX (REGC ) UNIV CALIFORNIA.  
 PI Chrispeels MJ, Colucci MG, Moore JG;  
 XX  
 DR WPI; 1999-081274/07.  
 DR N-PSDB; AAX03593.  
 XX  
 PT New nucleic acid encoding plant lectin that preserves progenitor  
 PT cells - particularly haematopoietic progenitors, useful for bone  
 PT marrow reconstitution after ablative therapy, and to increase DNA  
 PT transfer in gene therapy  
 XX  
 XX Claim 1; Page 30-31; 72pp; English.

XX The present sequence represents a lectin derived progenitor cell  
 CC preservation factor. The protein is used to preserve unipotent,  
 CC pluripotent or totipotent progenitor cells, especially haematopoietic  
 CC cells, and also progenitors from nerve, muscle, skin, gut, bone,  
 CC kidney, liver, pancreas or thymus. Specific applications are  
 CC preservation of cultured cells intended for administration after  
 CC (anticancer) myeloblastic therapy (bone marrow or whole-body irradiation  
 CC or chemotherapy) to reconstitute the haematopoietic system; enrichment  
 CC of progenitor cells (e.g. during ex vivo purging of malignant cells);  
 CC treatment of tissues containing haematopoietic progenitors for subsequent  
 CC transplant to improve haematopoietic competence; improving transfer of  
 CC exogenous DNA to progenitor cells (in gene therapy of various  
 CC haematological disorders, e.g. sickle-cell anaemia); and protection  
 CC against ablative therapy (to eliminate proliferating cells specifically),  
 CC followed by re-establishment of differentiation and proliferation of  
 CC preserved progenitors. The protein, when linked to magnetic beads, may  
 CC also be used to isolate cells that express the FLK2/FLT3 receptor.

XX  
 SQ Sequence 264 AA;

Query Match 65.9%; Score 796.5; DB 20; Length 264;  
 Best Local Similarity 66.7%; Pred. No. 2.1e-63;  
 Matches 156; Conservative 30; Mismatches 45; Indels 3; Gaps 2;

Qy 1 AQSVSFTTFKFDSDQKLMFQGHGTTSSNVLTQLKDSNGNPVSTSVGRVLYSAPLRLWE 60  
 Db 1 AQSLSFSFTTFEDPQEDLIFOGHATSTNNVLTQKDSAGNPVSSAGRVLYSAPLRLWE 60

Qy 61 SSTVSVSTFTTFOISTPYTSPGDLGAFFLAPYDVTVPNSAGNLLGLFPNLNLRNS 120  
 Db 61 DSAVLTSFDITINFEISTPYTSRIADGLAFTIAPPDSVI--SYHGGFLGLFPNANTLNS 118

Qy 121 TTSKETTIDVNAASNVAVEFDTPNDNIGDPYRKHGIDVNSIRSKATVAVDWMQNGKT 180  
 Db 119 STSENQQT-TTKAASSNVVAVEFDTPYLPDYPYHIGIDVNSIRSKATKWMQNGKI 177

Qy 181 ATAHISYNSASKRLSVTTFYPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234  
 Db 178 ATAHISYNSVSKRLSVTSYAGSKPATLSYDIELHTVLPENVRVGLSASTGQDK 231

RESULT 3  
 AAG62890  
 ID AAG62890 standard; Protein; 264 AA.  
 XX  
 AC AAG62890;  
 XX  
 DT 17-SEP-2001 (first entry)  
 XX  
 DE Amino acid sequence of a hyacinth bean FRIL polypeptide.  
 XX  
 KW FRIL; FLK2/Flt3 tyrosine kinase receptor-interacting lectin;  
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;  
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
 KW severe combined immunodeficiency; aplastic anemia; tissue repair.  
 XX  
 OS Dolichos lab lab.  
 XX

PN WO200149851-A1.  
 XX 12-JUL-2001.  
 XX 30-DEC-1999; 99WO-US31307.  
 XX 30-DEC-1999; 99WO-US31307.  
 XX (PHYL-) PHVLOGIX LLC.  
 XX Colucci MG, Chrispeels MJ, Moore JG;  
 DR WPI; 2001-441882/47.  
 DR N-PSDB; AAH42287.  
 XX Legume Progenitor cell preservation factors for in vivo or ex vivo  
 PT preservation of hematopoietic progenitor cells and as therapeutics for  
 PT alleviating/reducing progenitor cell-depleting activity of cancer  
 PT therapeutics -  
 XX  
 XX Example 1; Page 54-55; 173pp; English.  
 PS  
 CC The present sequence represents a FRIL (Flk2/Flt3 tyrosine kinase  
 CC receptor-interacting lectin) polypeptide. The specification describes a  
 CC composition of one or more members of FRIL family of progenitor cell  
 CC preservation factors. The composition is useful for alleviating or  
 CC reducing the hematopoietic progenitor cell-depleting activity of  
 CC a therapeutic treatment, including radiotherapeutic and/or  
 CC chemotherapeutic treatments. Administration of FRIL compositions to a  
 CC patient prior to treatment of the patient with a therapeutic treatment  
 CC having a hematopoietic progenitor cell-depleting activity alleviates or  
 CC reduces the hematopoietic progenitor cell-depleting activity of the  
 CC therapeutic treatment in the patient. FRIL family members are useful for  
 CC isolating population of progenitor cells, hemangioblasts, and mesenchymal  
 CC stem cells. The composition is administered to reduce progenitor cell  
 CC depleting effects of chemotherapeutics, so that the patient can receive  
 CC a higher dose of the chemotherapeutic and preferably recover from cancer.  
 CC It is also administered to patients having, or predisposed to developing  
 CC a condition where the patients hematopoietic progenitor cells are  
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
 CC The isolated mesenchymal cells are useful for tissue repair.  
 XX  
 SQ Sequence 264 AA;  
 Query Match 65.9%; Score 796.5; DB 22; Length 264;  
 Best Local Similarity 66.7%; Pred. No. 2.1e-65;  
 Matches 156; Conservative 30; Mismatches 45; Indels 3; Gaps 2;  
 QY 1 AQSVSFTTKFDSQKDLMFQGHITSSNVLTQLTKLDSNGNPVSTSVGRVLSAPRLWE 60  
 DB 1 AQSLSFSFTKFDPNQEDLIFQGHATSTNNVLQVTKLDSAGNPVSSAGRVLSAPRLWE 60  
 QY 61 SSTVSTFTFTTFOISNPTSPPGDGLAFFLAPYDVTIPPNSAGNLGLFNLNLRNS 120  
 DB 61 DSAVLTSFDTIINFEISTPTYSRIADGLAFFIAPPDSVI--SYHGFLGLFNLNLRNS 118  
 QY 121 TTSKETTIDVNAASNNVAVFEDTYPNDIGDPYRKHIGIDVNSIRSKATVADWQNGKT 180  
 DB 119 STSENQT--TTKAASNNVAVFEDTYPNDIGDPYRKHIGIDVNSIRSKATVADWQNGKI 177  
 QY 181 ATAHISYNSAKRLSVTTTYPGGKAVLSHVDVLTQVLPQWTRVGFSASTGLEK 234  
 DB 178 ATAHISYNSKRLSVTSYYAGSKPATLSYDIELHTVLPWVRVGLSASTGQDK 231  
 RESULT 4  
 AAG62894  
 ID AAG62894 standard; Protein; 286 AA.  
 XX  
 AC AAG62894;  
 XX  
 DT 17-SEP-2001 (first entry)  
 XX

DE Alpha-amylase inhibitor signal peptide and FRIL fusion.  
 XX  
 KW FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;  
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;  
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
 KW severe combined immunodeficiency; aplastic anemia; tissue repair;  
 KW alpha-amylase inhibitor gene.  
 XX  
 OS Synthetic.  
 OS Unidentified.  
 OS Dolichos lab lab.  
 PN WO200149851-A1.  
 XX 12-JUL-2001.  
 XX 30-DEC-1999; 99WO-US31307.  
 XX 30-DEC-1999; 99WO-US31307.  
 XX (PHYL-) PHVLOGIX LLC.  
 XX Colucci MG, Chrispeels MJ, Moore JG;  
 PI WPI; 2001-441882/47.  
 DR N-PSDB; AAH42295.  
 XX Legume Progenitor cell preservation factors for in vivo or ex vivo  
 PT preservation of hematopoietic progenitor cells and as therapeutics for  
 PT alleviating/reducing progenitor cell-depleting activity of cancer  
 PT therapeutics -  
 XX  
 XX Example 1; Page 59; 173pp; English.  
 PS  
 CC The present sequence represents fusion protein of alpha-amylase inhibitor  
 CC signal peptide and FRIL (Flk2/Flt3 tyrosine kinase receptor-interacting  
 CC lectin). The specification describes a composition of one or more members  
 CC of FRIL family of progenitor cell preservation factors. The composition  
 CC is useful for alleviating or reducing the hematopoietic progenitor  
 CC cell-depleting activity of a therapeutic treatment, including  
 CC radiotherapeutic and/or chemotherapeutic treatments. Administration of  
 CC FRIL compositions to a patient prior to treatment of the patient with  
 CC a therapeutic treatment having a hematopoietic progenitor cell-depleting  
 CC activity alleviates or reduces the hematopoietic progenitor  
 CC cell-depleting activity of the therapeutic treatment in the patient.  
 CC FRIL family members are useful for isolating population of progenitor  
 CC cells, hemangioblasts, and mesenchymal stem cells. The composition is  
 CC administered to reduce progenitor cell depleting effects of  
 CC chemotherapeutics, so that the patient can receive a higher dose of the  
 CC chemotherapeutic and preferably recover from cancer. It is also  
 CC administered to patients having, or predisposed to developing a  
 CC condition where the patients hematopoietic progenitor cells are  
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
 CC The isolated mesenchymal cells are useful for tissue repair.  
 XX  
 SQ Sequence 286 AA;  
 Query Match 65.9%; Score 796.5; DB 22; Length 286;  
 Best Local Similarity 66.7%; Pred. No. 2.4e-65;  
 Matches 156; Conservative 30; Mismatches 45; Indels 3; Gaps 2;  
 QY 1 AQSVSFTTKFDSQKDLMFQGHITSSNVLTQLTKLDSNGNPVSTSVGRVLSAPRLWE 60  
 DB 23 AQSLSFSFTKFDPNQEDLIFQGHATSTNNVLQVTKLDSAGNPVSSAGRVLSAPRLWE 82  
 QY 61 SSTVSTFTFTTFOISNPTSPPGDGLAFFLAPYDVTIPPNSAGNLGLFNLNLRNS 120  
 DB 83 DSAVLTSFDTIINFEISTPTYSRIADGLAFFIAPPDSVI--SYHGFLGLFNLNLRNS 140  
 QY 121 TTSKETTIDVNAASNNVAVFEDTYPNDIGDPYRKHIGIDVNSIRSKATVADWQNGKT 180  
 DB 141 STSENQT--TTKAASNNVAVFEDTYPNDIGDPYRKHIGIDVNSIRSKATVADWQNGKI 199

QY 181 ATAHISYNSASKRLSVTTFYPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234  
 ID AAG62898 standard; Protein: 303 AA.  
 Db 200 ATAHISYNSVKRLSVTSYAGSKPATLSYDIELHTVLPWVRVGLSASTGQDK 253

## RESULT 5

AC AAG62898;  
 DT 17-SEP-2001 (first entry)  
 XX  
 DE Amino acid sequence of a french bean FRIL polypeptide.  
 DE FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;  
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;  
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
 KW severe combined immunodeficiency; aplastic anemia; tissue repair.  
 XX  
 OS Phaseolus vulgaris.  
 XX  
 PN WO200149851-A1.  
 XX  
 PD 12-JUL-2001.  
 XX  
 PF 30-DEC-1999; 99WO-US31307.  
 XX  
 PR 30-DEC-1999; 99WO-US31307.  
 XX  
 PA (PHYL-) PHYLOGIX LLC.  
 XX  
 PI Colucci MG, Chrispeels MJ, Moore JG;  
 XX WPI; 2001-441882/47.  
 DR N-PSDB; AAH42306.  
 XX

PT Legume Progenitor cell preservation factors for in vivo or ex vivo  
 PT preservation of hematopoietic progenitor cells and as therapeutics for  
 PT alleviating/reducing progenitor cell-depleting activity of cancer  
 PT therapeutics -  
 XX  
 XX Example 5; Page 81; 173pp; English.  
 XX  
 CC The present sequence represents a FRIL (FLK2/FLT3 tyrosine kinase  
 CC receptor-interacting lectin) polypeptide. The specification describes a  
 CC composition of one or more members of FRIL family of progenitor cell  
 CC preservation factors. The composition is useful for alleviating or  
 CC reducing the hematopoietic progenitor cell-depleting activity of  
 CC a therapeutic treatment, including radiotherapeutic and/or  
 CC chemotherapeutic treatments. Administration of FRIL compositions to a  
 CC patient prior to treatment of the patient with a therapeutic treatment  
 CC having a hematopoietic progenitor cell-depleting activity alleviates or  
 CC reduces the hematopoietic progenitor cell-depleting activity of the  
 CC therapeutic treatment in the patient. FRIL family members are useful for  
 CC isolating population of progenitor cells, hemangioblasts, and mesenchymal  
 CC stem cells. The composition is administered to reduce progenitor cell  
 CC depleting effects of chemotherapeutics, so that the patient can receive  
 CC a higher dose of the chemotherapeutic and preferably recover from cancer.  
 CC It is also administered to patients having, or predisposed to developing  
 CC a condition where the patients hematopoietic progenitor cells are  
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
 CC The isolated mesenchymal cells are useful for tissue repair.  
 XX  
 SQ Sequence 303 AA;

Query Match 64.2%; Score 776.5; DB 22; Length 303;  
 Best Local Similarity 66.8%; Pred. No. 1.8e-63;  
 Matches 157; Conservative 25; Mismatches 48; Indels 5; Gaps 4;

QY • 1 AQSVSFTTFKDSQKDLMFQGH-TISSNVIOQLTKLDSNGNPVSTSVGRVLYSAPRLWE 60  
 Db 1 AQSLSFNFTKFDLQKDLIFQGDATSTNNVLQTLKDSGNGPVGASVGRVLYSAPRLWE 60

QY 61 SSTVSTFETFTFOISTPPTSPDGGLAFFLAPYDVTIPPNSAGNLGLFPLNALRNS 120  
 Db 61 NSMAVSSFETNLTIQISTPHYYAAGFAFFLAPDHTVIPPNSWGFGLGYSNV--FRNS 118  
 QY 121 TTSKETTI-DVNAASNNVAVFEFTYPPNDNIGDPYKHKHIGIDVNSIRSKATVAMDQNGK 179  
 Db 119 PTSENOQSGDVNTDS-RVVAVEFTFPNANI-DPNVRHIGIDVNSIKSKETARWEQNGK 176  
 QY 180 TATAHISYNSASKRLSVTTFYPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234  
 Db 177 TATARISYNSASKRSTVTTFYPMGVVALSHDVLHAEPLPEWVRVGLSASTGEEK 231

## RESULT 6

AA58736  
 ID AAY58736 standard; Protein: 290 AA.

AC AAY58736;  
 XX  
 DT 25-APR-2000 (first entry)  
 XX  
 DE Jack bean concanavalin A lectin.  
 XX  
 KW Concanavalin A; Con A; lectin; jack bean; insecticide;  
 KW transgenic plant; Brassica; insect resistance.  
 XX  
 OS Canavalia ensiformis.

FH Key Location/Qualifiers  
 FT Modified-site 152  
 FT /note= "N-glycosylated"

PN WO200001223-A1.

PD 13-JAN-2000.

XX 02-JUL-1999; 99WO-SE01209.

XX 03-JUL-1998; 98SE-0002425.

XX (PLAN-) PLANT SCI SVERIGE AB.

PI Ahman I, Melander M, Vamling K;

XX WPI; 2000-160693/14.

DR N-PSDB; AAZ58017.

XX Novel lectins used to produce transgenic Brassica plants which are  
 PT resistant to insects -

XX Example 2a; Fig 4; 51pp; English.

CC The present sequence is that of the concanavalin A (Con A) lectin  
 CC of jack bean. The present invention relates to: a transformed  
 CC Brassica plant that is resistant to certain insect pests; an  
 CC expression cassette containing DNA that codes for at least 1  
 CC lectin selected from Con A, modified Con A (see AAY58737) and pea  
 CC lectin (see AAY58738); transgenic plant cells containing at least 1  
 CC copy of the DNA; a new lectin derived from the jack bean Con A  
 CC gene; a method of imparting resistance to insects selected from  
 CC blossom beetles (pollen beetles) of the genus *Meligethes*, flea  
 CC beetles of the genus *Phyllotreta*, and root flies of the genus  
 CC *Delia*; and a method for protecting a plant against infestation by  
 CC insects of these genera.

XX Sequence 290 AA;

Query Match 53.2%; Score 643.5; DB 21; Length 290;  
 Best Local Similarity 54.0%; Pred. No. 3.3e-51;  
 Matches 127; Conservative 35; Mismatches 64; Indels 9; Gaps 4;

QY 3 SVSFTTFKDSQKDLMFQGH-TISSNVIOQLTKLDSNGNPVSTSVGRVLYSAPRLWE 61



```

Db 36 ALHEMFNQSKDQKDLILOGDATTGTDGNLELTRVSSNGSPQSGSVGRALFYAPVHIWES 95
QY 62 STVYSTETTTFTQISTPTSPPGDGLAFFAPYDTVTPPNSAGNLLGLFPNLALRNST 121
Db 96 SAVVASEATFTFLIKSP-DSHPADGIAFFTSNIDSSIPSGSTGRLLGLFPDANVIRNS- 153
QY 122 TSKETTTIDVNAASN--NVVAVEFDTPNDNIGDPYRKHIGIDVNSRKSATVADWONGK 179
Db 154 ----TTIDFNAAYNADTIVAVELDTYPNTDIDGDPSPHIGIDIKSVRSKKTAKWNQNGK 209
QY 180 TATAHISYNSASKRLSVTTFYPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234
Db 210 VGTAAHIINYNDKRLSAVSVYPNADSATVSDVDLDNVLPEWVRVGLSASTGLYK 264

RESULT 7
AAV58737
ID AAV58737 standard; Protein; 290 AA.
XX
AC AAV58737;
XX
DT 25-APR-2000 (first entry)
XX
XX Jack bean concanavalin A N152S mutant.
XX
KW Concanavalin A; Con A; lectin; jack bean; insecticide;
KW transgenic plant; Brassica, insect resistance; mutant; mutein.
XX
OS Canavalia ensiformis.
XX
FH Key Location/Qualifiers
FT misc_difference 152
FT /note= "replaces wild-type Asn"
XX
XX WO200001223-A1.
XX
XX 13-JAN-2000.
XX
XX 02-JUL-1999; 99WO-SE01209.
XX
XX 03-JUL-1998; 98SE-0002425.
XX
XX (PLAN-) PLANT SCI SVERIGE AB.
XX
XX Alman I, Melander M, Vamling K;
XX
XX WPI; 2000-160693/14.
XX
XX N-PSDB; AAZ58018.
XX
XX Novel lectins used to produce transgenic Brassica plants which are
XX resistant to insects
XX
XX Example 2a; Fig 6; 51pp; English.
XX
XX The present sequence is that of a jack bean concanavalin A
XX (Con A) mutein in which the Asn-152 residue of the native protein
XX is replaced by Ser in order to disrupt a glycosylation signal. The
XX present invention relates to: a transformed Brassica plant that is
XX resistant to certain insect pests; an expression cassette
XX containing DNA that codes for at least 1 lectin selected from Con A
XX (see AAY58736), modified Con A and pea lectin (see AAY58738); transgenic
XX plant cells containing at least 1 copy of the DNA; the present
XX lectin derived from the jack bean Con A gene; a method of imparting
XX resistance to insects selected from blossom beetles (pollen
XX beetles) of the genus Meligethes, flea beetles of the genus
XX Phyllotreta, and root flies of the genus Delia; and a method for
XX protecting a plant against infestation by insects of these genera.
XX
XX Sequence 290 AA;

```

Query Match 52.8%; Score 638.5; DB 21; Length 290;  
 Best Local Similarity 53.6%; Pred. No. 9.4e-51;

```

Matches 126; Conservative 36; Mismatches 64; Indels 9; Gaps 4;
QY 3 SVSFTTKFSDQKDLMFQGH-TISSNNVIOLTKLDSNCPVSTVSGRVLSAPRLWES 61
Db 36 ALHEMFNQSKDQKDLILOGDATTGTDGNLELTRVSSNGSPQSGSVGRALFYAPVHIWES 95
QY 62 STVYSTETTTFTQISTPTSPPGDGLAFFAPYDTVTPPNSAGNLLGLFPNLALRNST 121
Db 96 SAVVASEATFTFLIKSP-DSHPADGIAFFTSNIDSSIPSGSTGRLLGLFPDANVIRNS- 153
QY 122 TSKETTTIDVNAASN--NVVAVEFDTPNDNIGDPYRKHIGIDVNSRKSATVADWONGK 179
Db 154 ----TTIDFNAAYNADTIVAVELDTYPNTDIDGDPSPHIGIDIKSVRSKKTAKWNQNGK 209
QY 180 TATAHISYNSASKRLSVTTFYPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234
Db 210 VGTAAHIINYNDKRLSAVSVYPNADSATVSDVDLDNVLPEWVRVGLSASTGLYK 264

RESULT 8
AAV74765
ID AAR74765 standard; protein; 237 AA.
XX
AC AAR74765;
XX
DT 13-OCT-1995 (first entry)
XX
DE Legume concanavalin A.
XX
KW Pentraxin; legume; lectin; structural homology; serum amyloid P; SAP;
KW amyloid; fibril; human; Limulus; C-reactive protein; CRP; pea;
KW female hamster protein; 3-D structure; jellyroll topology;
KW picornavirus; coat protein; pentameric structure; concanavaline A.
XX
OS Legume sp..
XX
XX WO9505394-A.
XX
XX 23-FEB-1995.
XX
XX 17-AUG-1994; 94WO-GB01802.
XX
XX 17-AUG-1993; 93GB-0017120.
XX
XX (BIRK-) BIRKBECK COLLEGE.
XX
XX (ROYA-) ROYAL POSTGRAD MED SCHOOL.
XX
XX Blundell TL, Pepys MB;
XX
XX WPI; 1995-098720/13.
XX
XX New cpds. Inhibiting binding of serum amyloid P to amyloid
XX fibrils - produced by computer assisted molecular design, useful
XX for preventing, treating or diagnosing amyloidosis, e.g.
XX Alzheimer's disease
XX
XX Claim 7; Fig 4; 72pp; English.
XX
XX The sequences given in AAR74763-70 represent various pentraxins and
XX legume lectins. All these proteins show structural homology and may
XX be used in the molecular design of a molecule for the inhibition of
XX serum amyloid P (SAP) binding to amyloid fibrils. The similarities
XX in the amino acid sequences of SAP, human and Limulus C-reactive
XX protein (CRP) and female hamster protein suggests that they may have
XX similar 3-D structures. The jellyroll topology of the pentraxins is
XX reminiscent of the picornavirus coat proteins which also have
XX pentameric structures. However, pentaxins resemble more closely
XX legume lectins such as concanavaline A and pea lectin. Alignment
XX of sequences on the basis of topologically equivalent features of the
XX three dimensional structures shows that helices occupy different
XX positions in the pentraxins and legume lectins and that the amino acid
XX sequences of the two families have identities of only approx. 11%. The
XX two main helices in SAP occur before and after strand L, whereas the

```

CC helices in the legume lectins occur at the C-terminus of strand J.  
 CC There is a long insertion between the end of the helix after strand D  
 CC and the beginning of strand E in the lectins relative to pentraxins.  
 CC Strands G, H and I together with the type IV beta-hairpin between H and  
 CC I are identical in both SAP and pea lectin. The so-called pentraxin  
 CC octapeptide signature sequence, HXCXS/TWXS, is in this region so it is  
 CC not conserved in the legume lectins.  
 XX  
 SQ Sequence 237 AA;  
 Query Match 44.5%; Score 537.5; DB 16; Length 237;  
 Best Local Similarity 49.3%; Pred. No. 1.5e-41;  
 Matches 104; Conservative 30; Mismatches 58; Indels 19; Gaps 2;  
 QY 24 TISSNNVIQLTKLDSNGPNVSTSGVRLVLSAPLRLWESSTVSTFTTFQISTPTSP 83  
 Db 1 TTGTCNLELTRVSSNGPSSGSGVGRALFPVPHIVHESAAATVSTFEATFLIKSP-DSH 59  
 QY 84 PDGGLAFFLAPDYTVIPPNASAGNLLGLFPLNALRNSSTTSKETIDVNAASNNVAVVEFD 143  
 Db 60 PADGTAFITSNIDSSIPSGSTGRILGLFPDANA-----DTIVAVELD 101  
 QY 144 TYPNDNIGDPYRKHHIGIDVNSIRSKATVAMDWONGKTATATAHISVNSASKRLSVTTEYPGG 203  
 Db 102 TYPNDIGDPSPHIGIDIKSVRSKKTAKNMQDGKVGTAHHIYNSVDKRLSAVVSPNA 161  
 QY 204 KAVLSLHSHVELTQVLPQWIRVGFSASTGLEK 234  
 Db 162 DATSVSYDVLNDVLPQWIRVGLSASTGLYK 192  
 RESULT 9  
 AAY06812  
 ID AAY06812 standard; protein; 285 AA.  
 AC AAY06812;  
 XX  
 XX 23-JUN-1999 (first entry)  
 DE R. pseudoacacia lectin RPBAl polypeptide subunit b.  
 KW Robinia pseudoacacia; lectin; medicine; mucosal cell proliferation;  
 KW cancer; metabolic disorder; mucositis; cytotoxic.  
 XX  
 OS Robinia pseudoacacia.  
 XX  
 PN WO9911278-A1.  
 PD 11-MAR-1999.  
 PF 28-AUG-1998; 98WO-GB02612.  
 PR 29-AUG-1997; 97GB-0018413.  
 PA (ALIZ-) ALIZYME THERAPEUTICS LTD.  
 XX  
 PI Bardocz ZM, Fish NW, Palmer RMJ, Pusztai AJ;  
 DR WPI; 1999-228935/19.  
 XX  
 PT Use of Robinia pseudoacacia lectin in medicine  
 XX  
 PS Disclosure: Fig 1; 37pp; English.  
 CC The invention relates to Robinia pseudoacacia lectins which can be used  
 CC in medicine. The lectin is used for the control of mucosal cell  
 CC proliferation, for the reduction and/or treatment of damage caused by  
 CC cell damaging agents, especially in the treatment of cancer, and/or for  
 CC the reduction and/or treatment of metabolic disorders. It is especially  
 CC useful for the treatment of mucositis in mammalian cells and/or tissues,  
 CC particularly human cells and/or tissues (especially mucous cells  
 CC including mucous membrane). Use of the lectin is effective and does not  
 CC damage the gut as is the case with cytotoxic drugs and radiation. The

CC present sequence represents a polypeptide subunit of R. pseudoacacia  
 CC lectin RPBAl.  
 XX  
 SQ Sequence 285 AA;  
 Query Match 37.8%; Score 457; DB 20; Length 285;  
 Best Local Similarity 44.7%; Pred. No. 5.3e-34;  
 Matches 105; Conservative 40; Mismatches 68; Indels 22; Gaps 8;  
 QY 3 SVSFTTTFKFDSDOKLDFQGHHT-ISSNNVIQLTKLDSNGPNVSTSGVRLVLSAPLRLWES 61  
 Db 34 SLSSFPFKKISQDLIFQSDALVTSKGVQLTTV--NDGRVYDSIGRVLYIAAPFQIWD 91  
 QY 62 ST-VVSTETFTTFOISTPTSPPGDGLAFFLAPDYTVIPPNASAGNLLGLFPLNALRNS 120  
 Db 92 TTGNVASPVTSEFIIKAPNEGKTADGLVFLAPVGST-QPLKGGGLGLF-----KDE 144  
 QY 121 TTSKETIDVNAASNNVAVBEDTYPNDNIGDPYRKHHIGIDVNSIRSKATVAMDWONGKT 180  
 Db 145 SYNK-----SNOIVAVEFDIFRN-VAMPNGIHMGIDVNSIQSVRTVRWDWANGEV 194  
 QY 181 ATAHTSYNSASKRLSVTTFYPG-GKAVLSLHSHVELTQVLPQWIRVGFSASTGLEK 234  
 Db 195 ANVFLSYEASTKSLTASLVPSLEKSFILSAIVDLKKVLPQWIRVGFATTGLSE 249  
 RESULT 10  
 AAY06811  
 ID AAY06811 standard; protein; 285 AA.  
 AC AAY06811;  
 XX  
 XX 23-JUN-1999 (first entry)  
 DE R. pseudoacacia lectin RPBAl polypeptide subunit a.  
 KW Robinia pseudoacacia; lectin; medicine; mucosal cell proliferation;  
 KW cancer; metabolic disorder; mucositis; cytotoxic.  
 XX  
 OS Robinia pseudoacacia.  
 XX  
 PN WO9911278-A1.  
 PD 11-MAR-1999.  
 PF 28-AUG-1998; 98WO-GB02612.  
 PR 29-AUG-1997; 97GB-0018413.  
 PA (ALIZ-) ALIZYME THERAPEUTICS LTD.  
 XX  
 PI Bardocz ZM, Fish NW, Palmer RMJ, Pusztai AJ;  
 DR WPI; 1999-228935/19.  
 XX  
 PT Use of Robinia pseudoacacia lectin in medicine  
 XX  
 PS Disclosure: Fig 1; 37pp; English.  
 CC The invention relates to Robinia pseudoacacia lectins which can be used  
 CC in medicine. The lectin is used for the control of mucosal cell  
 CC proliferation, for the reduction and/or treatment of damage caused by  
 CC cell damaging agents, especially in the treatment of cancer, and/or for  
 CC the reduction and/or treatment of metabolic disorders. It is especially  
 CC useful for the treatment of mucositis in mammalian cells and/or tissues,  
 CC particularly human cells and/or tissues (especially mucous cells  
 CC including mucous membrane). Use of the lectin is effective and does not  
 CC damage the gut as is the case with cytotoxic drugs and radiation. The  
 CC present sequence represents a polypeptide subunit of R. pseudoacacia  
 CC lectin RPBAl.  
 XX  
 SQ Sequence 285 AA;

Query Match 36.9%; Score 446; DB 20; Length 285;  
Best Local Similarity 42.6%; Pred. No. 5.5e-33;  
Matches 100; Conservative 47; Mismatches 66; Indels 2

[illegible]

RESULT 11  
AAY06814  
ID AAY06814 standard; protein; 285 AA.

XX	AY06814;
XX	AC
XX	DT
XX	23-JUN-1999 (first entry)
DE	Amino acid sequence of lectin RPAI from <i>R. pseudoacacia</i> seed.
XX	
KW	<i>Robinia pseudoacacia</i> ; lectin; medicine; mucosal cell proliferation
KW	cancer; metabolic disorder; mucositis; cytotoxic.

XX	Robinia pseudoacacia.	
OS		
XX		
PN	WO9911278-A1.	
XX		
XX	11-MAR-1999.	
PD		
XX		
XX	28-AUG-1998; 98WO-GB02612.	
XX		
PF		
XX		
PR	29-AUG-1997; 97GB-0018413.	
XX		
XX	(ALIZ-) ALIZYME THERAPEUTICS LTD.	
PA		
XX		
PI	Bardocz ZM, Fish NW, Palmer RMJ, Pusztaí AJ;	
XX		
XX	WPI; 1999-228935/19.	
DR		
XX		
XX		
PT	Use of Robinia pseudoacacia lectin in medicine	
XX		
PS	Disclosure; Fig 2; 37pp; English.	
XX		
CC	The invention relates to Robinia pseudoacacia lectins which can be used	
CC	in medicine. The lectin is used for the control of mucosal cell	
CC	proliferation, for the reduction and/or treatment of damage caused by	
CC	cell damaging agents, especially in the treatment of cancer, and/or for	
CC	the reduction and/or treatment of metabolic disorders. It is especially	
CC	useful for the treatment of mucositis in mammalian cells and/or tissues,	
CC	particularly human cells and/or tissues (especially mucous cells	
CC	including mucous membrane). Use of the lectin is effective and does not	
CC	damage the gut as is the case with cytotoxic drugs and radiation. The	
CC	present sequence represents a R. pseudoacacia lectin RPA1. This lectin	
CC	is composed of two subunits- polypeptide a and polypeptide b	
CC	(AAY06811-12).	
XX		
SQ	Sequence 285 AA;	

Query Match 36.8%; Score 445; DB 20; Length 285;  
Best Local Similarity 41.9%; Pred. No. 6.8e-33;  
Matches 98; Conservative 49; Mismatches 65; Indels 22; Gaps 8;

Qy	3
Db	34

Qy	62	ST-VVSTETTTTPTPOISTPYTPSPGGDLAFLFAPYDTVIPNSAGNLLGLFPNNALRNS	120
Dd	93	TGTVNAVSEVTFSFIQAIPNPATTADGLAFLFAPVDI--QPGLGGMLGIF-----KDG	144
Qy	121	TTSKETITDVNAASNNVVAVEDTYTPNDNLGDYPKRKHIGIDVNSIRSKATVADWQNCKT	180
Dd	145	SYNK-----SNQIVAVEEDTFENIHFDPKGRHMGINVSIVSVKTKPPWNNTGEV	194
Qy	181	ATAHISYSASKKRLSMTTFYPGGKAVLSLHD-VELTQVLQPIWRGVGFSASTGLE	233
Dd	195	ANVFISEASTKSLNASLYPSLETFSFIHAIVDVKDVLPEWRFSGSATGID	248

RESULT 12  
AAY06815  
ID AAY06815 standard; protein: 285 AA.

XX	AY06815;
AC	23-JUN-1999 (first entry)
XX	Amino acid sequence of lectin RpbII from <i>R. pseudoacacia</i> seed.
DT	<i>Robinia pseudoacacia</i> ; lectin; medicine; mucosal cell proliferation;
DT	cancer; metabolic disorder; mucositis; cytotoxic.
XX	
DE	
XX	
KW	
KW	

XX	
OS	<i>Robinia pseudoacacia</i> .
XX	
XX	
PN	WO9911278-A1.
XX	
XX	
PD	11-MAR-1999.

XX  
28-AUG-1998: 98WO-GB02612.

XX 29-AUG-1997; 97GB-0018413.

XX  
PA (ALIZ-) ALIZYME THERAPEUTICS LTD.

XX PI Bardocz ZM, Fish NW, Palmer RMJ, Pusztai AJ;

XX  
DR WPI; 1999-228935/19.

XX  
PT  
Use of Robina pseudoacacia lectin in medicine

XX  
PS Disclosure; Fig 2; 37pp; English.

XX The invention relates to Robinia pseudoacacia lectins which can be used  
CC in medicine. The lectin is used for the control of mucosal cell  
CC proliferation, for the reduction and/or treatment of damage caused by  
CC cell damaging agents, especially in the treatment of cancer, and/or for  
CC the reduction and/or treatment of metabolic disorders. It is especially  
CC useful for the treatment of mucositis in mammalian cells and/or tissues,  
CC particularly human cells and/or tissues (especially mucous cells  
CC including mucous membrane). Use of the lectin is effective and does not  
CC damage the gut as is the case with cytotoxic drugs and radiation. The  
CC present sequence represents a R. pseudoacacia lectin RPA11. This lectin  
CC is composed of the subunit- polypeptide c (AAV06813).  
XX  
SO Sequence 285 AA:

Sequence 285 AA;

Query Match  
36.7%; Score 444; DB 20; Length 285;

Best Local Similarity 42.7%; Pred. No. 8.5e-33;  
Matches 100; Conservative 46; Mismatches 66

3.  $\text{CVCCTTCTKCTNSDQVNI MEQCHT - TSSNNVTOI MKT DSCNNDVCTCVCQVNI VCAPI BIFMFS 61}$

QY 34 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

[illegible]

Db 93 TTNVASFVTSFIQAPNPATAGLAFAPVDT--QPLDGGMLGIF-----KNG 144  
 QY 121 TTSKETTIDVNAASNNVAVVEFDTPNDNIGDPYRKHIGIDVNSIRSKATVAMWQNGKT 180  
 Db 145 YFNK-----SNQIVAVEFDTFESNRH-WDPTGRHMGINVNSIVSVKTPWPNWANGCV 194  
 QY 181 ATAHISYNSASKRLSVTTTPPGKAVLSLHD-VELTQVLPQWIRVGFSASTGLE 233  
 Db 195 ANVFISYEASTKSLTASLYPSLETSLIIHAIIVDKVDLPENWVRFGFSATTGID 248

## RESULT 13

AA58738  
 ID AAY58738 standard; Protein; 275 AA.

AC AAY58738;

DT 25-APR-2000 (first entry)

XX Pea lectin.

DE Lectin; pea; insecticide; transgenic plant; Brassica;  
 KW insect resistance.

XX Pisum sativum.

XX WO200001223-A1.

XX 13-JAN-2000.

XX 02-JUL-1999; 99WO-SE01209.

XX 03-JUL-1998; 98SE-0002425.

XX (PLAN-) PLANT SCI SVERIGE AB.

XX Ahman I, Melander M, Vamling K;

XX WPI; 2000-160693/14.

XX N-PSDB; AAZ58019.

XX Novel lectins used to produce transgenic Brassica plants which are  
 resistant to insects

XX Example 2b; Fig 9; 51pp; English.

XX The present sequence is that of pea cv. Lincoln lectin. The  
 invention relates to: a transformed Brassica plant that is  
 resistant to certain insect pests; an expression cassette  
 containing DNA that codes for at least 1 lectin selected from Con A  
 (see AAY58736), modified Con A (see AAY58737) and pea lectin;  
 CC transgenic plant cells containing at least 1 copy of the DNA; a new  
 CC lectin derived from the jack bean Con A gene; a method of imparting  
 CC resistance to insects selected from blossom beetles (pollen  
 CC beetles) of the genus *Meligethes*, flea beetles of the genus  
 CC *Phyllotreta*, and root flies of the genus *Delia*; and a method for  
 CC protecting a plant against infestation by insects of these genera.

XX Sequence 275 AA;

Query Match 34.0%; Score 411.5; DB 21; Length 275;  
 Best Local Similarity 40.2%; Pred. No. 8.1e-30;  
 Matches 9; Conservative 39; Mismatches 72; Indels 33; Gaps 8;

QY 2 QSVSFTEKFDSDOKLDFQGHITSSSNVQLTKLDSNGNPVSTSVGRVLYSAPLRLWES 61  
 Db 32 ETTSELTIKFSPDQNLIFQGDGYTTKEKLTGK-----AVKNTVGRALYSSPHIHDR 85  
 QY 62 ST-VVSTFTFTFTFTQISTPYTSPGCDGLAFFLAPYDVTIPPNSAGNLLGLFPNLNLR 120  
 Db 86 ETGNVANEVTSFTFVINAPNSYNVADGTFETFTAPVDT--KPQTGGGYLGVE-----NS 136

QY 121 TTSKETTIDVNAASNNVAVVEFDTPNDNIGDPYR--KHIGIDVNSIRSKATVAMWQNG 178  
 Db 137 AEYDKTT-----QTVAVEFDTFYN-AAWDPNSNRDRHIGIDVNSIKSVNTKSKWLONG 187  
 QY 179 KTATAHISYNSASKRLSVTTTPPGG-----KAVLSLHDVELTQVLPQWIRVGFSASTGL 232  
 Db 188 EAVNVIAFNAATNVLTVSLTPYNSLEENVTSTYTLSDVSLKDVVPENWVRFGFSATTGA 247  
 QY 233 E 233  
 Db 248 E 248

## RESULT 14

AA45911  
 ID AAR45911 standard; protein; 242 AA.

AC AAR45911;

DT 18-JUL-1994 (first entry)

XX Bauhinia purpurea lectin larvicidal deriv.

DE Bauhinia purpurea lectin larvicidal deriv.  
 XX Non-lysine; insect larvae; sorghum; wheat; oat; rye; rice; European;  
 KW corn borer.

XX Bauhinia purpurea.

XX WO9402514-A.

XX 03-FEB-1994.

XX 22-JUL-1993; 93WO-US06946.

XX 24-JUL-1992; 92US-0921179.

XX (PION-) PIONEER HI-BRED INT INC.

XX Balasubramaniam NK, Rao AG;

XX WPI; 1994-048798/06.

XX Bauhinia purpurea lectin larvicidal non-lysine derivs. - and  
 PT corresp. DNA used to transform e.g. maize plants, to protect  
 PT plants or stored prods. from insect larvae, e.g. European corn  
 PT borer

XX Claim 1; Page 17; 19pp; English.

XX The sequence is that of a Bauhinia purpurea lectin which has had  
 CC lysine residues replaced by other amino acids which either preserve  
 CC the positive charge at the position of the substitution or provide a  
 CC neutral residue. The protein derivs. are effective larvicides  
 CC against insect pupae, such as the European corn borer. The protein  
 CC can be administered enterally to the larvae in their diet and can be  
 CC used to protect sorghum, wheat, oats, rye, etc.  
 CC See also AAR45912.

XX Sequence 242 AA;

Query Match 31.3%; Score 379; DB 15; Length 242;  
 Best Local Similarity 37.0%; Pred. No. 6.8e-27;  
 Matches 88; Conservative 47; Mismatches 77; Indels 26; Gaps 7;

QY 6 FTFTKFDSDOK----DLMFOGHITSSSNVQLTKLDSNGNPVSTSVGRVLYSAPLRLWES 61  
 Db 8 FTFPNFWNTQENGTELIIFLGNTATYTPGALRLTRIGDGIPLKSNAGQASYSRPVFLWDS 67  
 QY 62 STVVSFTFTFTF---QISTPYTSPGCDGLAFFLAPYDVTIPPNSAGNLLGLFPNLNLR 118  
 Db 68 TGHVASFYTSFIVRSIDVPHIT--ADGFAFLAPDVSSV--KDYGGCLGLF-----R 117  
 QY 119 NSTTSKETTIDVNAASNNVAVVEFDTPNDNIGDPYRKHIGIDVNSIRSKATVAMWQNG 178

```

Db 118 YKTA-----DSKNQVAVEDTWTENSDLRYPHIGINVSATVTRNDNDA 170
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 179 --KTATAHISNSAKRLSVTFYPGGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 171 YVTKSTAHITYDATSKIITVLLTYDNGRHYQLSHVVDLPKILPERVIRIGSGGTGFNE 228
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
AAI27265
ID AAY27265 standard; protein; 262 AA.
XX
AC AAY27265;
XX
DT 22-OCT-1999 (first entry)
XX
DE B. purpurea native lectin sequence.
XX
KW Bauhinia purpurea lectin; BPL; larvicide; insect control;
KW European corn borer; Ostrinia nubilalis; crop protection;
KW insect larvae infestation.
XX
OS Bauhinia purpurea.
XX
PN US5945589-A.
XX
PD 31-AUG-1999.
XX
PF 24-MAR-1993; 93US-0038761.
XX
PR 24-MAR-1993; 93US-0038761.
XX
PT 24-JUL-1992; 92US-0921179.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
Balasubramaniam NK, Rao AG;
WPI; 1999-508204/42.

```

Derivatives of Bauhinia purpurea lectins useful as larvicides to control insects (e.g. European corn borer) in field crops such as corn, wheat and rapeseed

Example 5; Columns 11-14; 8pp; English.

The invention describes compounds derived from Bauhinia purpurea lectin ((BPL), a 262 amino acid peptide comprising 7 lysine residues - the present sequence) by the replacement of one or more lysine residues with other amino acid residues which either preserve the positive charge or provide a neutral residue at the point of substitution. The compounds are useful as larvicides against insects such as the European corn borer (Ostrinia nubilalis). The compounds may be used to protect commercial and field crops from insect larvae infestation, especially by the European corn borer. The BPL-derivatives are effective against insects which may be resistant to the naturally occurring form of BPL. Crops which may be protected using the BPL derivatives are corn/maize (Zea mays), rye, barley, wheat (Triticum aestivum), sorghum (Sorghum bicolor), oats, millet, rice, triticale, sunflower, alfalfa, rapeseed and soybean), fiber crops (such as cotton), fruit crops (such as melons) and vegetable crops (such as onion, pepper, tomato, cucumber, squash, carrot, crucifer (e.g. cabbage, broccoli and cauliflower), eggplant, spinach, potato and lettuce). The compounds are useful for protecting species from the genera Fragaria, Lotus, Medicago, Onobrychis, Trifolium, Trigonella, Vigna, Citrus, Linum, Geranium, Manicot, Daucus, Arabidopsis, Brassica, Raphanus, Sinapis, Atropa, Capsicum, Datura, Hyoscyamus, Lycopersicon, Nicotiana, Solanum, Petunia, Digitalis, Majorana, Cichorium, Helianthus, Lactuca, Bromus, Asparagus, Antirrhinum, Hemerocallis, Nemesis, Pelargonium, Panicum, Pennisetum, Ranunculus, Senecio, Salpiglossis, Cucumis, Browallia, Glycine, Lolium, Triticum and Datura. The present sequence represents the B. purpurea native lectin sequence.

Sequence 262 AA;

```

Query Match 31.3%; Score 379; DB 20; Length 262;
Best Local Similarity 37.0%; Pred. No. 7.6e-27;
Matches 88; Conservative 47; Mismatches 77; Indels 26; Gaps 7;

QY 6 FTFTKFDSDOK----DLMFQHTRISSNVQLTKLDSNGNPVSTSVGRVLSAPLRLWES 61
      ||| || : : : : : : : : : : : : : : : : : : : : : :
Db 8 FTFPNWSNTQENGTEIFLGNATYTPGALRLTRIGEDGIPKSNAGQASYRPPVFLWDS 67
      : : : : : : : : : : : : : : : : : : : : : :

QY 62 STVVSTFETTFP---QISTPYTSPPGDGLAFFLAPYDTVIPPNAGNLLGLFPNLNLR 118
      : : : : : : : : : : : : : : : : : : : : : :
Db 68 TGHVASFTSFSFIVRSIDVPHIT--ADGFAFLAPVDSSV--RDYGGCLGLF-----R 117
      : : : : : : : : : : : : : : : : : : : : : :

QY 119 NSTTSKETTIDVNAASNNVAVEDTYPNDNIGDPYRKKHIGIDVNSIRSKATVAMDWONG 178
      : : : : : : : : : : : : : : : : : : : : : :
Db 118 YKTA-----DPSKNQVAVEDTWTENSDLRYPHIGINVSATVTRNDNDA 170
      : : : : : : : : : : : : : : : : : : : : : :

QY 179 --KTATAHISNSAKRLSVTFYPGGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234
      : : : : : : : : : : : : : : : : : : : : : :
Db 171 YVTKSTAHITYDATSKIITVLLTYDNGRHYQLSHVVDLPKILPERVIRIGSGGTGFNE 228
      : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: February 26, 2003, 16:43:20  
Job time : 29.6292 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:40:33 ; Search time 8.17978 Seconds  
(without alignments)  
1079.114 Million cell updates/sec

Title: US-09-476-485A-8

Perfect score: 1209

Sequence: 1 AQSVSFTFTKFDSDQKDLMF.....TQVLPOWIRGFSASTGLEK 234

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174566 seqs, 37721826 residues

Total number of hits satisfying chosen parameters: 174566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	249	20.6	669	10	US-09-862-027-27
2	101	8.4	2383	10	US-09-912-020-302
3	97	8.0	2344	10	US-09-815-242-12713
4	94.5	7.8	345	10	US-09-813-820-7
5	90.5	7.5	1169	10	US-09-801-368-106
6	84.5	7.0	448	10	US-09-871-212-8
7	83.5	6.9	888	10	US-09-826-752-6
8	83.5	6.9	1723	10	US-09-841-132-394
9	83.5	6.9	1723	10	US-09-841-132-395
10	83.5	6.9	5795	10	US-09-815-242-12610
11	83	6.9	283	10	US-09-925-299-984
12	83	6.9	595	10	US-09-881-752A-166
13	81.5	6.7	599	9	US-09-771-382-6
14	81.5	6.7	599	10	US-09-787-862-15
15	80.5	6.7	1367	10	US-09-801-368-108
16	80.5	6.7	2092	9	US-10-147-026-12
17	79.5	6.6	507	9	US-09-738-626-5458
18	79.5	6.6	922	10	US-09-886-468-19
19	79	6.5	906	10	US-09-905-983-46

20	79	6.5	906	10	US-09-746-491-46	Sequence 46, Appl
21	79	6.5	2026	10	US-09-801-368-86	Sequence 86, Appl
22	79	6.5	2224	9	US-10-115-563-14	Sequence 14, Appl
23	78.5	6.5	1848	9	US-09-839-996-6	Sequence 6, Appl
24	78	6.5	413	10	US-09-746-491-43	Sequence 43, Appl
25	77	6.4	422	9	US-09-738-626-5947	Sequence 5947, Ap
26	77	6.4	969	9	US-09-981-353-122	Sequence 122, App
27	77	6.4	977	10	US-09-925-297-797	Sequence 797, App
28	76.5	6.3	632	10	US-09-853-533A-8	Sequence 8, Appl
29	76.5	6.3	1447	10	US-09-797-097-2	Sequence 2, Appl
30	76	6.3	579	10	US-09-870-203A-6	Sequence 6, Appl
31	76	6.3	579	10	US-09-870-203A-8	Sequence 8, Appl
32	76	6.3	579	10	US-09-870-203A-18	Sequence 18, Appl
33	76	6.3	579	10	US-09-870-203A-10	Sequence 20, Appl
34	76	6.3	580	10	US-09-870-203A-12	Sequence 12, Appl
35	76	6.3	580	10	US-09-870-203A-10	Sequence 10, Appl
36	76	6.3	581	10	US-09-758-008-5	Sequence 5, Appl
37	76	6.3	581	10	US-09-870-203A-2	Sequence 2, Appl
38	76	6.3	581	10	US-09-870-203A-4	Sequence 4, Appl
39	76	6.3	581	10	US-09-870-203A-14	Sequence 14, Appl
40	76	6.3	581	10	US-09-870-203A-16	Sequence 16, Appl
41	76	6.3	600	10	US-09-871-212-5	Sequence 5, Appl
42	75.5	6.2	451	9	US-10-074-279-6	Sequence 6, Appl
43	75.5	6.2	594	10	US-09-815-108-22	Sequence 22, Appl
44	75.5	6.2	660	9	US-09-872-462-4	Sequence 4, Appl
45	75.5	6.2	704	10	US-09-801-368-218	Sequence 218, App

#### ALIGNMENTS

##### RESULT 1

US-09-862-027-27

; Sequence 27, Application US/09862027

; Patent No. US20020142428A1

; GENERAL INFORMATION:

; APPLICANT: Hodge, Martin R.

; TITLE OF INVENTION: NO. US20020142428A1el Kinases and Uses Thereof

; CURRENT APPLICATION NUMBER: US/09/862,027

; CURRENT FILING DATE: 2001-05-21

; PRIOR APPLICATION NUMBER: US 09/345,473

; PRIOR FILING DATE: 1999-06-30

; NUMBER OF SEQ ID NOS: 82

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 27

; LENGTH: 669

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

US-09-862-027-27

Query Match

Best Local Similarity 20.6%; Score 249; DB 10; Length 669;

Mismatches 79; Conservative 41; Mismatches 81; Indels 40; Gaps 10;

QY	13	SDQKDLMEFGHTISSNNVQLTKLDSNG-----NPVSTSVGRVLYSAPRLWES-STVVS	66
DB	23	SEEGEGFNGVLYDSNG-IAIT--NSKGLMKLTNSEFSYGHVFNPSVFRNPNGTVS	79
QY	67	TFETTFQISTPYTSPPGDGLAFFLAPYDVTPPNAGNLLGLFNLNLRNNTTSKET	126
DB	80	SFSTTFVFAIVSNVNALDGHGLAFVISP-TKGLPYSSSSQYLCLF-----NLNNGD-	130
QY	127	TIDVNAASNNVAVVEFDYPNDNIGDPYRKHGIDVNSIRS-KAIVAWDQ-----	176
DB	131	-----PSNHIVAVEFDIFQNOEFDMDNNHVGIDINLSSEKASTAGYEDDGTFKNI	184
QY	177	---NCKTATAHISYNSAKRLSVT---TFYGGKAVSLSHDVELTQVLQWIRVGFSA	230
DB	185	RLINQKPIQAWIEWDSSRQLNVTIHPILPKPIPLLSLTLDSPYLFDSMIVGFTSAT	244
QY	231	G 231	

Db 245 G 245

## RESULT 2

US-09-912-020-302  
; Sequence 302, Application US/09912020  
; Patent No. US20020045592A1  
; GENERAL INFORMATION:  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Trawick, John  
; APPLICANT: Forsyth, R. Allyn  
; APPLICANT: Froelich, Jamie M.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN

; FILE REFERENCE: ELITRA.001DV1

; CURRENT APPLICATION NUMBER: US/09/912,020

; CURRENT FILING DATE: 2001-07-23

; PRIOR APPLICATION NUMBER: 09/492,709

; PRIOR FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: 60/117,405

; PRIOR FILING DATE: 1999-01-27

; NUMBER OF SEQ ID NOS: 485

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 302

; LENGTH: 2383

; TYPE: PRT

; ORGANISM: E. Coli

US-09-912-020-302

## Query Match

Best Local Similarity 8.4%; Score 101; DB 10; Length 2383;

Matches 64; Conservative 37; Mismatches 101; Indels 96; Gaps 11;

QY 13 SOKDLMFQGHIT-----SSNVQLTKLDSNGNPVS----- 44

Db 1257 SDNKTVHFIGDFAAAKIELAPVPSIIAGTPQNSGSGVITATVVDNNGSPVKGVTNFT 1316

QY 45 -----TSVGR-----VLVSAPRLWESSTVSTFTFTTFTQISTPTVTPPGD 86

Db 1317 SNAETAEMTNGQAVNEOGKATVYTNTRSSIESCARPTVASELNGSSTLSTSNVN 1376

QY 87 G-----LAFFLAPYDVTIPPSAGNLLGLFPNL-NALRNSSTSKETIDVNA-----S 134

Db 1377 ADASTAHLTLQALEDV-----SAGETTSLYIEVKDNYGNGVPOQEVTLSPSPSEGVTPS 1432

QY 135 NNVAVEFTYDNIQDPYRKHIGIDVNSIRSKATVAMDWQNGKTATAHISY----- 187

Db 1433 NNAI-----YTNHNDGNFYASFTATKAGVYQLTATL-----ENGDSMQQVTVYVPNVANA 1482

QY 188 -----NSAKRLSVTTFYPGKAVSLGSHDVELTQVLPQWIRVGFSASTG 231

Db 1483 EITLAASKDPVADNNDLTLTATVADTEGNAIA---NTEVTTLPEDEVKRNATLSDG 1537

## RESULT 3

US-09-815-242-12713

; Sequence 12713, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; TITLE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12713

; LENGTH: 2344

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-815-242-12713

Query Match 8.0%; Score 97; DB 10; Length 2344;

Best Local Similarity 24.2%; Pred. No. 3.2;

Matches 58; Conservative 33; Mismatches 85; Indels 64; Gaps 12;

QY 2 QSVFTFTFKFSDQKDLMFQGHIT---SSNVQLTKLDSNGNPVSTSVGRV----- 50

Db 565 QSVYYFT--DVKAPTIVGNGQTIEVGKTMNPVLTITDNGTGTVTNTVGLPSGLSYDS 622

QY 51 ----LYSAPRLWESS-TVSTFE-----TFTTFOI---STPYTSPGDLGLAFLAPYD 96

Db 623 ATNSIIGTPTKIGSTVTVSTDOANNKSTTTFTTNVDTTAPTVTPIGDKSSEVFSPIS 682

QY 97 TVIPNSAGNLLGLFPNLNALRNSITTSKETTIDVNAANNVV-----AVEFDTPNDNIGD 152

Db 683 PI-----NATQDNG-----NAVNTVGLPSGLFDDSTNTTIST 719

QY 153 PYRKHIGIDVNSIRSKATVAMDWQNGKTATA---HISYNSASKRLSVTTFYPGKAVSL 209

Db 720 P--INIGTSTITI-----VSTDAGNKTTTTFKYEVTRNMSDSVSTSGTQSQSVST 772

## RESULT 4

US-09-813-820-7

; Sequence 7, Application US/09813820

; Patent No. US20020102262A1

; GENERAL INFORMATION:

; APPLICANT: Hook, Magnus

; APPLICANT: Patti, Joseph M.

; APPLICANT: House-Pompeo, Karen

; APPLICANT: Sthanam, Narayana

; APPLICANT: Symersky, Jindrich

; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS

; TITLE OF INVENTION: AND METHODS OF USE

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSSEE: Arnold, White &amp; Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: Texas

; COUNTRY: U.S.

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/813,820



```
; FILING DATE: 22-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/856,253
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: TANK:193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 345 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-813-820-7

Query Match          7.8%; Score 94.5; DB 10; Length 345;
Best Local Similarity 24.18; Pred. No. 0.41;
Matches 59; Conservative 32; Mismatches 73; Indels 81; Gaps 16;

QY 3 SVSFTFTKFDSDQKDLMPQGHITSSNNVIQLTKLDSNGN-PVSTSGRVL-----51
Db 107 NVIYFTDYVNTKDDVK---ATLMPAIDPENVKKTGNVLATIGISTANKTVLVDYE 163
QY 52 -YSAPLRWESSTVVSFTFTFTQISTPYSPGDDGLAFLAPYDVTIPPNSAGNLLGL 110
Db 164 KYGFYLSIKGTIDQIDKNTNTYR-QTIYVNPSC-----DNVIAPVLTGNLK-- 210
QY 111 FPNL--NAL--RNSTTSKETTIDVNA-----SNNWAVE 141
Db 211 -PNTDSNALIDQQTISKVKYVD-NARDLSSEYFVNPENFEDVNSVITPPNHOYKVE 268
QY 142 FDTYPNDNIGDPYKHHG--IDVNS-----KATVAV---DWQNGKTATAHIS 186
Db 269 FNT-PDDQITTPYIVVYVNGHIDPNKSGDLALRSTLYGYNSNIWRMSWDN-----EVA 321
QY 187 YNSAS 191
Db 322 FNGS 326

RESULT 5
US-09-801-368-106
; Sequence 106, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US-03-07
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 106
; LENGTH: 1169
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-106

Query Match          7.5%; Score 90.5; DB 10; Length 1169;
Best Local Similarity 21.3%; Pred. No. 4.9;
Matches 55; Conservative 32; Mismatches 112; Indels 59; Gaps 9;

QY 2 QSVSFTFTKFDSDQKDLMPQGHITSSNNV-IQLTKLDSNGNPVSTSGRVLVSAPLRW 60
Db 498 ESTSYVTPYVSSSTAAANYTSSFSSEVCTETES-----TSTSTPYVTSS---WS 549
QY 61 SSTVSTFTFTFTQISTPYSPGDDGLAFLAPYDVTIPPNSAGNLLGLFPNLALRNS 120
Db 550 SSEVCTETETESYVTPYVS-----SSTAAANYTSSFSSEVCTE 592
QY 121 TTSKEIT-----IDVNAASNNVVA--VEFTYPNDNIGDPYKHHGIDVNS 164
Db 593 CTETESTSTPYATSTGATSTASTNTMTSLVQDITVVSFSLSTVSEHTNAPTSS 652
QY 165 IRSKATVAMWQNGKTATAHISYNSASKRLSVTTFYPGGKAVSLSHDYEL-----214
Db 653 VESNASTFISNKGSKV-----SYVTSSIH-SITPMYPSNOTVTSSSVVSTPITSESS 707
QY 215 --TQVLQWIRVGFSA 230
Db 708 ASVTILPSTITSEFKPST 725

RESULT 6
US-09-871-212-8
; Sequence 8, Application US/09871212
; Patent No. US20020034519A1
; GENERAL INFORMATION:
; APPLICANT: Tikoo, Suresh
; APPLICANT: Babluk, Lorne
; APPLICANT: Zhang, Linong
; APPLICANT: Wu, Qiaohua
; TITLE OF INVENTION: MODIFIED BOVINE ADENOVIRUS HAVING
; FILE REFERENCE: 293102003000
; CURRENT APPLICATION NUMBER: US/09/871,212
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/208,678
; PRIOR FILING DATE: 2000-05-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Porcine adenovirus 3
US-09-871-212-8

Query Match          7.0%; Score 84.5; DB 10; Length 448;
Best Local Similarity 25.2%; Pred. No. 4.8;
Matches 36; Conservative 23; Mismatches 63; Indels 21; Gaps 6;

QY 18 LMFQGHITSSNNVIQLTKLDSNGNPVSTSGRVLVSAPLRWESSTVSTFTFTFTQIS 77
Db 294 LVLGSAFALKSS--IDLTSMTKKVNFIDGAGRL-----QSD---SYKGFGRFSN 339
QY 78 TPTTSPPGDGLA-FFLAPYDVTIPPNSAGNLLGLFPNLALRNSTTSKETTIDVNAASN 136
Db 340 DSVIEPTAAGLSPAWLMPSTFIYPRNTSGSLTSFVYIN-----QTVVHVDIKVNTLSTN 394
QY 137 VVAVEFTYPNDNIGDPYKHHG 159
Db 395 GYSLEEN-FQNMFSFSAFFSTSYG 416
```



```

RESULT 11
US-09-925-299-984
; Sequence 984, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925, 299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124, 270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 984
; LENGTH: 283

```

379 NLMAOKDAFI.NVHKNTQ'

379 NLMAOKDAFLNVHKNTOTAVAOAOETYPBPSVINTNNYCOMV-----CVD-----AMACV 127

```

QY      174  DWONGKT-----ATAHISYNSAS 191
      |      :      :      :      :
      |      :      :      :      :
Ddb     428  KWFFGKTKRFGSYGYSYNHAN 451

RESULT 13
US-09-771-382-6
; Sequence 6, Application US/09771382
; Patent No. US20020160016A1
; GENERAL INFORMATION:
; APPLICANT: Peak, Ian
; APPLICANT: Jennings, Michael
; TITLE OF INVENTION: MODIFIED SURFACE ANTIGEN
; FILE REFERENCE: 8795-24U1
; CURRENT APPLICATION NUMBER: US/09/771,382
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: US 60/177,917

```

```

; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-771-382-6

Query Match          6.7%; Score 81.5; DB 9; Length 599;
Best Local Similarity 18.8%; Pred. No. 13;
Matches 40; Conservative 43; Mismatches 75; Indels 55; Gaps 10;

QY 12 DSDQKLMFOGHTTSSNNVLTQLKLDNGNPVSTSVGRV-----LYSAPLRL-- 58
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 53 DEDEEEEL--EPVYRSALVLOF-MIDKEGGENESTGNIGWSIYYDNINTLHGATVTLKA 109
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 59 -----WESSWTSFTFTTFQIS---TPVTSPPGDGLAFFLAPYDIVIPENSAGN 106
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 110 GDNLKI KONTNKNNTNENTNDSFYSLKKDLTDLTSVETKLSF-----GANGN 158
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 107 LLGLFPLNAL-----RNSTTSKETITDVNAASNVAVEFDT-----YPNDNIGDPYK 157
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 159 KVNITSDTKGLNFAKETAGTNGDTVHLNGIGSTLDTLLNTGATNTVNDNVTDKKRR 218
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 158 IGIDVNSTRSKATVAWDMQN---GKATAHISY 187
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 219 AA-----SVKDVNLNAGWNKGYKPGTTASDNVDF 247
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
US-09-797-862-15
; Sequence 15, Application US/09797862
; Patent No. US2002010276A1

```

```

: APPLICANT: PEAK, IAN RICHARD ANSELM
: APPLICANT: JENNINGS, MICHAEL PAUL
: APPLICANT: MOXON, E, RICHARD
: TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
: FILE REFERENCE: 065064/0134
: CURRENT APPLICATION NUMBER: US/09/797.862
: CURRENT FILING DATE: 2001-05-03
: PRIOR APPLICATION NUMBER: PCT/AU98/01031
: PRIOR FILING DATE: 1998-12-14
: PRIOR APPLICATION NUMBER: GB 9726398.2
: PRIOR FILING DATE: 1997-12-12
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 15
: LENGTH: 599
: TYPE: PRT
: ORGANISM: Neisseria meningitidis
: US-09-797-862-15

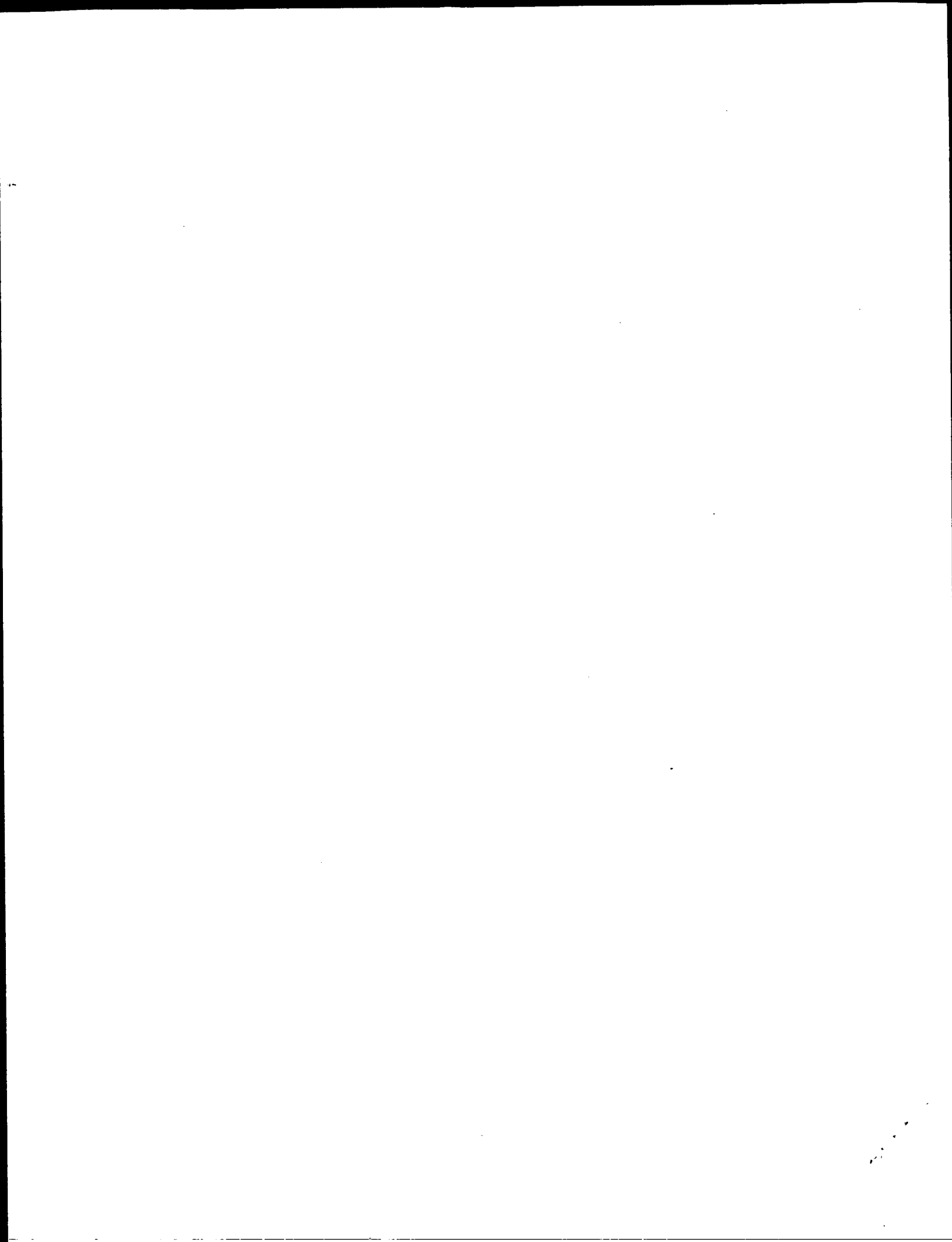
```

Wed Feb. 26 17:04:16 2003

us-09-476-485a-8.rapb

Page 7

Search completed: February 26, 2003, 16:52:43  
Job time : 17.1798 secs









[illegible]

```
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/881.189B
FILING DATE: June 24, 1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: 381-44 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 132 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-881-189B-12

Query Match 25.1%; Score 402; DB 4; Length 132;
Best Local Similarity 65.6%; Pred. No. 1.2e-33;
Matches 80; Conservative 15; Mismatches 25; Indels 2; Gaps 2;

QY 132 DSRVAVVEFDT-FPNANI-DPNYRHIGIDVNSIKSKETARWQNGKTATARIYSNSASK 189
:|||||: : ||| |||||: ||| ||||| ||| ||||| ||| ||||| ||
Db 3 ESNVAVVEFDTDLNPDGDPNYIHIGIDVNSIRSVTASDWQNGKIATAHISYNSVK 62
:|||||: : ||| |||||: ||| ||||| ||| ||||| ||| ||||| ||
QY 190 KSTVTFYPGMEVVALSHDVLHAEPLPEWVRVGLSASTGEEKOKNTIISWFTSSLKNE 249
:|||||: : ||| |||||: ||| ||||| ||| ||||| ||| ||||| ||
Db 63 RLSVTYTPGRGKPKATSYDIELHTVLPWVRVGLSASTGQNIERTVHWSFTSSLTNV 122
:|||||: : ||| |||||: ||| ||||| ||| ||||| ||| ||||| ||
QY 250 VK 251
:
Db 123 AK 124

RESULT 8
US-08-038-761A-1
; Sequence 1, Application US/08038761A
; Patent No. 5945589
; GENERAL INFORMATION:
; APPLICANT: Rao, A. Gururaj; Kumar, M. Arun
; TITLE OF INVENTION: Derivatives of Bauhinia
; TITLE OF INVENTION: purpurea Lectin and Their Use as Larvicides
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pioneer Hi-Bred International, Inc.
; STREET: 700 Capital Square, 400 Locust
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft WORKS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/038,761A
; FILING DATE: 24-JUL-1992
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Roth, Michael J.
```

```
REGISTRATION NUMBER: 29,342
REFERENCE/DOCKET NUMBER: 0210 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 245-3594
TELEFAX: (515) 245-3634
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 262 residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION:
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: Bauhinia purpurea
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-08-038-761A-1

Query Match 24.0%; Score 383; DB 2; Length 262;
Best Local Similarity 40.3%; Pred. No. 3.1e-31;
Matches 96; Conservative 39; Mismatches 85; Indels 18; Gaps 7;

QY 17 DLIFQDATSTNNVLQTLKLDGNNPVCASVGRVLFSAFFHLWNSMAVSFETNLTIQI 76
:|||||: : ||| |||||: ||| ||||| ||| ||||| ||| ||||| ||
Db 23 EIIFLGNATYTPGALRLTRIGEDGIPKSNAGQASYSRPVFLWDSTGHVASFTSFSFIV 82
:|||||: : ||| |||||: ||| ||||| ||| ||||| ||| ||||| ||
QY 77 -STPHYYAAGDGAFFLAPHTVIPPNSWGRKFLGLYSNVFRNSPTSENOSFGDVNTDSRV 135
:|||||: : ||| |||||: ||| ||||| ||| ||||| ||| ||||| ||
Db 83 RSIDVPHITADGFAFFLAPVDSSV--KDYGGCLGLFRYKATATDP-SKNQ-----V 129
:|||||: : ||| |||||: ||| ||||| ||| ||||| ||| ||||| ||
QY 136 VAVEFDTPNAN-IDPNYRHIGIDVNSIKSKETARWQNG--KTATARIYSNSASKST 192
:|||||: : ||| |||||: ||| ||||| ||| ||||| ||| ||||| ||
Db 130 VAVEFDTPNTEWSDLRYPHIGINNVSVVATWDDNDAYVTKSTAHTIYDATSKIIT 189
:|||||: : ||| |||||: ||| ||||| ||| ||||| ||| ||||| ||
QY 193 VTTYPGMEVVALSHDVLHAEPLPEWVRVGLSASTGEEKOKNTIISWFTSSLKNEV 250
:|||||: : ||| |||||: ||| ||||| ||| ||||| ||| ||||| ||
Db 190 VLLTYDNGRHYQLSHVDLPKILPERVIRGPGSGTG-FNETQYILSWSFSTLNSTKI 246
:|||||: : ||| |||||: ||| ||||| ||| ||||| ||| ||||| ||
```

```

; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-881-189B-13

Query Match      16.7%; Score 267.5; DB 4; Length 105;
Best Local Similarity 55.8%; Pred. No. 5.3e-20;
Matches        63; Conservative    14; Mismatches   23; Indels     13; Gaps       4;

QY 1 AQSLSFNFTKFDLDOKDLIFOGDATSTNNVLQLTKLDSGPNVPGASVGRVLFSAPFLWE 60
Db 1 AQSLSFSTKFKDPNQEDLIFOGTATS-----KIDSAGNPVS SAGRVLXSAPRLWE 52
QY 61 NSMAYSSSETNLTIQISTPHPIYAADGRAFFLAHPDHTVIPPSNWKCKFLGLYSN 113
Db 53 DSAVLTSFDP--TIFYNTYSRIADGLA-FIAPPDSVI--SYHGGFGLGFPN 100

RESULT 11
US-09-228-986-77
; Sequence 77, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 77
; LENGTH: 632
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-77

Query Match      13.9%; Score 221.5; DB 4; Length 632;
Best Local Similarity 27.9%; Pred. No. 5e-14;
Matches        84; Conservative    50; Mismatches   106; Indels     61; Gaps       14;

QY 6 FNFTKFDLDOKDLIFOGDAT--STNNVLQLTKLDSGPNVPGASVGRVLFSAPFLWENSM 63
Db 40 ENGTF-----NLILEANASVIGSESVLSLT-----NHSHEFMLGRALYPAAVQM-KNH 86
QY 64 AVSSEFINLTIQISTPHPIYAADGRAFFLAHPDHTVIPPSNWKCKFLGLYSNVFRNSPTSEN 123
Db 87 TVSSFSTTFVFSPVPSPNEGGHLAFIMTYTPMGAKP-VQYLGLL-----NLTSNG 139
QY 124 QSGDVNTDSRVAVAEFDTEPFNANI-DPNRYRHIGIDVNISIKS---KETARWEW----- 172
Db 140 QPY-----NHLFAVEFDITINVEFKOPDRHVGVNDINSIVQTETAGY-WNGEFHEL 192
QY 173 --ONGKTATARSYNASKSKSTVTTFYPM----EVVALSHDVDLHAELPEWVRVGSLAST 227
Db 193 NLRSRNIQAVIDYDHLESSLNVTITVAGLP RPQRPLISLQIDLQNIVEEKMLGFSAAT 252
QY 228 GEKKQNTIIISWETS-----SLKNNEVKEPKEDMYIANVRSVTWTINDVLSY 275
Db 253 GLLVEDHYILAWSFTEDTAPLDVCSLSPANNYSPLSRGTAGV-----TVSVVLFW 308

; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-881-189B-13

Query Match      19.3%; Score 308.5; DB 3; Length 274;
Best Local Similarity 34.7%; Pred. No. 1.5e-23;
Matches        86; Conservative    40; Mismatches   97; Indels     25; Gaps       9;

QY 2 QSLSFNFTKFDLDOKDLIFOGDAT-STNNVLQLTKLDSGPNVPGASVGRVLFSAPFLWE 60
Db 31 EGISFNFTNETRGDQGVTLIQANIMANGILALT--NHTNPTW-NTRGALSXPVPLWD 86
QY 61 NSMA-VSSEFINLTIQISTPHPIYAADGRAFFLAHPDHTVIPPSNWKCKFLGLYSNVFRNSP 119
Db 87 SATGNVASFTVSFVVKEKGIPADGI VFFLAP-EARI PDNSAGGOLGT----- 136
QY 120 TSENSQSGDVNTDSRVAVAEFDTEPFNANIDPNRYRHIGIDVNISIKSETARWONGKTAT 179
Db 137 VNANKAY-----NPPFVGVEFDTSN-NWPDKSAHIGIDASLLISLRKVKNKYGSLVK 189
QY 180 ARTSYNASKSSTVTTPGMEVVVALSHVDLHAELPEWVRVGLSA--STGEEKQKNTII 237
Db 190 VSIIYDSLSTLSLVVTHENGQISTIAQQVVDLKAVLGKVRVGTAATTGTRYVELYDIH 249
QY 238 SWGSTSL 245
Db 250 AWSFTSL 257

RESULT 10
US-08-881-189B-13
; Sequence 13, Application US/08881189B
; Patent No. 6310195
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; OPERATING SYSTEM: MS-DOS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,189B
; FILING DATE: June 24, 1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
```

QY 276 I 276  
 Db 309 L 309  
 RESULT 12  
 US-08-913-942-15  
 ; Sequence 15, Application US/08913942  
 ; Patent No. 6200578  
 ; GENERAL INFORMATION:  
 ; APPLICANT: St. Geme, Joseph  
 ; APPLICANT: Barenkamp, Stephen J.  
 ; TITLE OF INVENTION: HAEMOPHILUS ADHESION PROTEINS  
 ; NUMBER OF SEQUENCES: 19  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Flehr Hobbach Test Albritton & Herbert LLP  
 ; STREET: Four Embarcadero Center, Suite 3400  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: United States  
 ; ZIP: 94111-4187  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/913,942  
 ; FILING DATE: 29-DEC-1997  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/409,995  
 ; FILING DATE: 24-MAR-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US96/4031  
 ; FILING DATE: 22-MAR-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Vance, Dolly A.  
 ; REGISTRATION NUMBER: 39,054  
 ; REFERENCE/DOCKET NUMBER: A-61053-1/RFT/RMS/DAV  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 781-1989  
 ; TELEFAX: (415) 398-3249  
 ; TELEX: 910 277299  
 ; INFORMATION FOR SEQ ID NO: 15:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 679 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: unknown  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: protein  
 ; US-08-913-942-15

Query Match 6.7%; Score 107; DB 4; Length 679;  
 Best Local Similarity 23.8%; Pred. No. 0.033;  
 Matches 73; Conservative 48; Mismatches 120; Indels 66; Gaps 14;  
 QY 22 GDAITNNVLQTLKDSGPNVGASVGRVLSFAPFLW-----ENSMVSSFFETNLTIQ 75  
 Db 82 GEQGTTEVINL-NTDSSGNAVGSST--ITFKAGDNLKIKQSGNDFTYSLKELKNLT-S 137  
 QY 76 ISTPHPYAAGDGAFFLAPHDTVIPPNWSGKFLGLYSNVFNRN---SPTSENQSGF----- 127  
 Db 138 VETKLSFGANGKVDITSDANGKLAKTGNNGQNSNVHLNGIASTLDTLAGGTGTHV 197  
 QY 128 DVNTDS----RVVAVEFDTPFNANIDPNRYHIGIDVNSIKSKETARWQNGKTATARIS 183  
 Db 198 DTNIDAVNYHRAASQVDVLNSGWNIGQN---GNNVDFVRTYDTV--DFVNGANANVSVT 251  
 QY 184 YNSASKSKVTTFYPCMEVVALSHD-----VDLHAELPEWVRVGLSASTG 228  
 Db 252 ADTAHKTTTIRVDVTGLPVQVYVTEGKTVVKVGNYYKAKDDGSDMNQKVGELAKT- 310

QY 229 EEKOKNTIISWSFTSSLKNNNEVKPEKEDMYIANVVRSYTWINDVLSYISNKMVDALNNHH 288  
 Db 311 ----KVLVSASGTPVKISNVADGTED-----TDAVSF---KOLKALQD-- 348  
 QY 289 KYRCST 295  
 Db 349 KQVTLST 355  
 RESULT 13  
 US-09-268-347-26  
 ; Sequence 26, Application US/09268347  
 ; Patent No. 6335182  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Loosmore, Sheena M.  
 ; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS  
 ; FILE REFERENCE: 1038-860  
 ; CURRENT APPLICATION NUMBER: US/09/268,347  
 ; CURRENT FILING DATE: 1999-03-16  
 ; NUMBER OF SEQ ID NOS: 54  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 26  
 ; LENGTH: 679  
 ; TYPE: PRT  
 ; ORGANISM: Haemophilus influenzae  
 ; US-09-268-347-26

Query Match 6.7%; Score 107; DB 4; Length 679;  
 Best Local Similarity 23.8%; Pred. No. 0.033;  
 Matches 73; Conservative 48; Mismatches 120; Indels 66; Gaps 14;  
 QY 22 GDAITNNVLQTLKDSGPNVGASVGRVLSFAPFLW-----ENSMVSSFFETNLTIQ 75  
 Db 82 GEQGTTEVINL-NTDSSGNAVGSST--ITFKAGDNLKIKQSGNDFTYSLKELKNLT-S 137  
 QY 76 ISTPHPYAAGDGAFFLAPHDTVIPPNWSGKFLGLYSNVFNRN---SPTSENQSGF----- 127  
 Db 138 VETKLSFGANGKVDITSDANGKLAKTGNNGQNSNVHLNGIASTLDTLAGGTGTHV 197  
 QY 128 DVNTDS----RVVAVEFDTPFNANIDPNRYHIGIDVNSIKSKETARWQNGKTATARIS 183  
 Db 198 DTNIDAVNYHRAASQVDVLNSGWNIGQN---GNNVDFVRTYDTV--DFVNGANANVSVT 251  
 QY 184 YNSASKSKVTTFYPCMEVVALSHD-----VDLHAELPEWVRVGLSASTG 228  
 Db 252 ADTAHKTTTIRVDVTGLPVQVYVTEGKTVVKVGNYYKAKDDGSDMNQKVGELAKT- 310  
 QY 229 EEKOKNTIISWSFTSSLKNNNEVKPEKEDMYIANVVRSYTWINDVLSYISNKMVDALNNHH 288  
 Db 311 ----KVLVSASGTPVKISNVADGTED-----TDAVSF---KOLKALQD-- 348  
 QY 289 KYRCST 295  
 Db 349 KQVTLST 355

RESULT 14  
 US-09-210-361-4  
 ; Sequence 4, Application US/09210361  
 ; Patent No. 6284479  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Nichols, Scott E.  
 ; TITLE OF INVENTION: Substitutes for Modified Starches and  
 ; FILE REFERENCE: 0357CR  
 ; CURRENT APPLICATION NUMBER: US/09/210,361  
 ; CURRENT FILING DATE: 1998-12-11  
 ; EARLIER APPLICATION NUMBER: 09/007,999  
 ; EARLIER FILING DATE: 1998-01-16  
 ; EARLIER APPLICATION NUMBER: 08/478,704  
 ; EARLIER FILING DATE: 1995-06-07  
 ; EARLIER APPLICATION NUMBER: 09/009,620

EARLIER FILING DATE: 1998-01-20  
EARLIER APPLICATION NUMBER: 08/485,243  
EARLIER FILING DATE: 1995-06-07  
EARLIER APPLICATION NUMBER: 09/008,172  
EARLIER FILING DATE: 1998-01-16  
EARLIER APPLICATION NUMBER: 08/482,711  
EARLIER FILING DATE: 1995-06-07  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 1375  
TYPE: PRT  
ORGANISM: streptococcus mutans  
US-09-210-361-4

Query Match  
Best Local Similarity 6.3%; Score 101; DB 4; Length 1375;  
Matches 70; Conservative 42; Mismatches 104; Indels 80; Gaps 17;  
QY 48 GRVLSAPFHLWENSMVSSFETNLTIQISTP-----HPYYAAD-----GFAPFLA--- 93  
Db 385 GALLYS-----NNSKLTQANSNYRILNRTPTNQTGKKDPRYTADRTIGGYEFLLANDV 438  
QY 94 -PHDTVIPPN--SWGKFLGLYSNVRNSPTSENGSFGDVNTDS-RVVAVEEDTFPNAID 149  
Db 439 DNSNPVVQAEQLNWLHFLMFGNIYANDP-----DANFDSIRVDAVD-----NVD 483  
QY 150 PNYRHIGIDVNSIKSKETARWEQNGKTATARIYSNSA-SKKSTVTTFYPGMEVV----- 203  
Db 484 ADLLQTAGDY-----LKAAGKIHKNDKAANDHLSILEAWSYNDTPPYLHDDGDNINMDNR 538  
QY 204 -----ALSHDVLHAEPEWVRVGLSASTGEEKOKNTIISWFTSSLKKNVEKPEKED 256  
Db 539 LRLSLYSLAKPLNQSGMPLITNSLVNRTDDNAETAAPVSYSFIRA-HDSEVQD----- 593  
QY 257 MYIANVRSYTWIN-DVLSY-----ISNKMVDALNNHKKYVRCSTCMLF 299  
Db 594 -LIRNIIR--TEINPNVVGYSFTTEIKKAFIYNK--DLLATEKKYTHYNTALSY 644

## RESULT 15

US-09-740-274-4  
Sequence 4, Application US/09740274  
Patent No. 6465203  
GENERAL INFORMATION:  
APPLICANT: Nichols, Scott E.  
TITLE OF INVENTION: Glucan-containing Compositions and Paper  
FILE REFERENCE: 0357CRD  
CURRENT APPLICATION NUMBER: US/09/740,274  
PRIOR FILING DATE: 2000-12-19  
PRIOR APPLICATION NUMBER: 09/210,361  
PRIOR FILING DATE: 1998-12-11  
PRIOR APPLICATION NUMBER: 09/007,999  
PRIOR FILING DATE: 1998-01-16  
PRIOR APPLICATION NUMBER: 08/478,704  
PRIOR FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: 09/009,620  
PRIOR FILING DATE: 1998-01-20  
PRIOR APPLICATION NUMBER: 08/485,243  
PRIOR FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: 09/008,172  
PRIOR FILING DATE: 1998-01-16  
PRIOR APPLICATION NUMBER: 08/482,711  
PRIOR FILING DATE: 1995-06-07  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 1375  
TYPE: PRT  
ORGANISM: streptococcus mutans  
US-09-740-274-4

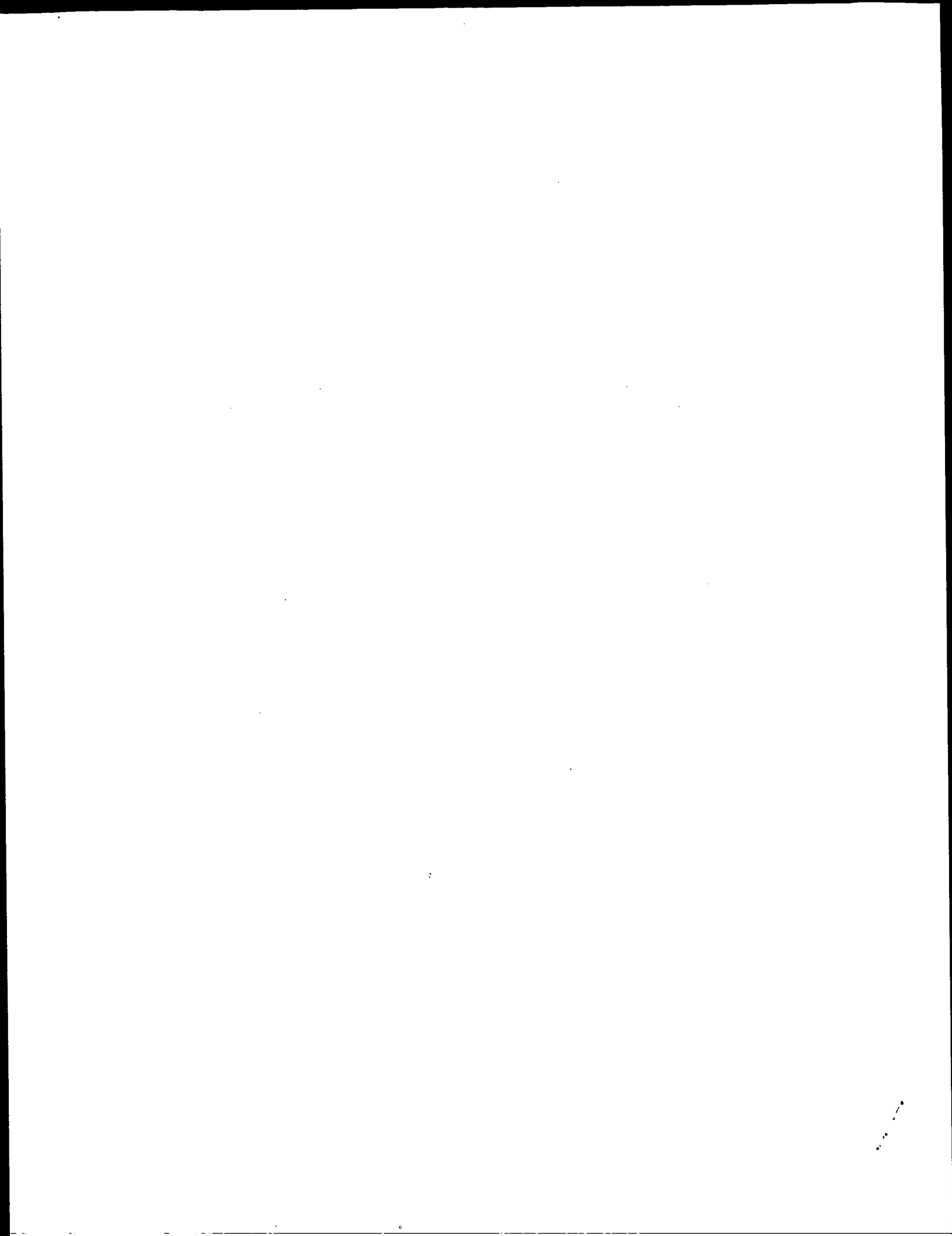
Query Match

6.3%; Score 101; DB 4; Length 1375;

Best Local Similarity 23.6%; Pred. No. 0.42;  
Matches 70; Conservative 42; Mismatches 104; Indels 80; Gaps 17;  
QY 48 GRVLSAPFHLWENSMVSSFETNLTIQISTP-----HPYYAAD-----GFAPFLA--- 93  
Db 385 GALLYS-----NNSKLTQANSNYRILNRTPTNQTGKKDPRYTADRTIGGYEFLLANDV 438  
QY 94 -PHDTVIPPN--SWGKFLGLYSNVRNSPTSENGSFGDVNTDS-RVVAVEEDTFPNAID 149  
Db 439 DNSNPVVQAEQLNWLHFLMFGNIYANDP-----DANFDSIRVDAVD-----NVD 483  
QY 150 PNYRHIGIDVNSIKSKETARWEQNGKTATARIYSNSA-SKKSTVTTFYPGMEVV----- 203  
Db 484 ADLLQTAGDY-----LKAAGKIHKNDKAANDHLSILEAWSYNDTPPYLHDDGDNINMDNR 538  
QY 204 -----ALSHDVLHAEPEWVRVGLSASTGEEKOKNTIISWFTSSLKKNVEKPEKED 256  
Db 539 LRLSLYSLAKPLNQSGMPLITNSLVNRTDDNAETAAPVSYSFIRA-HDSEVQD----- 593  
QY 257 MYIANVRSYTWIN-DVLSY-----ISNKMVDALNNHKKYVRCSTCMLF 299  
Db 594 -LIRNIIR--TEINPNVVGYSFTTEIKKAFIYNK--DLLATEKKYTHYNTALSY 644

Search completed: February 26, 2003, 16:41:25

Job time : 17.5094 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: February 26, 2003, 16:40:33 : Search time 10.5918 Seconds  
(without alignments)  
1079.114 Million cell updates/sec

Title: US-09-476-485A-6  
Perfect score: 1599  
Sequence: 1 AOSLSFNFTKFDLDQKDLIF.....LNNHKYVRCTCMLFMKKK 303

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 174566 seqs, 37721826 residues

Total number of hits satisfying chosen parameters: 174566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	202.5	12.7	669	10	US-09-862-027-27
2	101	6.3	1375	10	US-09-740-274-4
3	100	6.3	881	10	US-09-850-351A-32
4	97	6.1	3092	10	US-09-801-368-172
5	96.5	6.0	784	10	US-09-967-805-7
6	92.5	5.8	1024	10	US-09-879-257A-30
7	91.5	5.7	730	10	US-09-801-368-126
8	91	5.7	1250	10	US-09-801-368-364
9	89	5.6	425	10	US-09-850-351A-27
10	88	5.5	2042	9	US-10-192-584-6
11	87.5	5.5	234	10	US-09-815-242-5763
12	87.5	5.5	268	10	US-09-815-242-12678
13	87	5.4	632	10	US-09-853-533A-8
14	86.5	5.4	2383	10	US-09-912-020-302
15	85.5	5.3	589	10	US-09-972-086-2
16	85.5	5.3	1475	10	US-09-740-274-2
17	84	5.3	1066	9	US-09-423-126-3
18	84	5.3	1066	10	US-09-280-197-5
19	83.5	5.2	1198	10	US-09-866-582-36

20	83.5	5.2	1241	12	US-10-001-215-5	Sequence 5, Appli
21	82.5	5.2	552	10	US-09-909-464A-2	Sequence 2, Appli
22	82.5	5.2	555	9	US-10-121-032-24	Sequence 24, Appli
23	82.5	5.2	2076	10	US-09-815-242-5815	Sequence 5815, Ap
24	82.5	5.2	2186	10	US-09-815-242-12913	Sequence 12913, A
25	82.5	5.2	3014	10	US-09-737-149-2	Sequence 2, Appli
26	82.5	5.2	3034	10	US-09-737-149-25	Sequence 25, Appli
27	82.5	5.2	3034	10	US-09-737-149-30	Sequence 30, Appli
28	82	5.1	792	9	US-10-055-364-42	Sequence 42, Appli
29	81	5.1	1316	9	US-10-120-544A-4	Sequence 4, Appli
30	81	5.1	5795	10	US-09-815-242-12610	Sequence 12610, A
31	80.5	5.0	473	9	US-09-991-053-2	Sequence 2, Appli
32	80.5	5.0	589	10	US-09-866-510-12	Sequence 12, Appli
33	80.5	5.0	1089	9	US-09-955-363-36	Sequence 36, Appli
34	80.5	5.0	1089	10	US-09-769-987-2	Sequence 2, Appli
35	80.5	5.0	1089	10	US-09-919-497-90	Sequence 2, Appli
36	80.5	5.0	1089	10	US-09-866-510-2	Sequence 2, Appli
37	80.5	5.0	1089	10	US-09-866-510-4	Sequence 2, Appli
38	80.5	5.0	1089	10	US-09-866-510-6	Sequence 4, Appli
39	80.5	5.0	1089	10	US-09-866-510-8	Sequence 6, Appli
40	80.5	5.0	1089	10	US-09-866-510-10	Sequence 8, Appli
41	80.5	5.0	2001	9	US-10-072-621-8	Sequence 10, Appli
42	80	5.0	378	10	US-09-815-242-13445	Sequence 8, Appli
43	79.5	5.0	518	9	US-09-976-297-2	Sequence 13446, A
44	79.5	5.0	630	10	US-09-801-368-220	Sequence 2, Appli
45	79.5	5.0	716	10	US-09-853-533A-2	Sequence 2, Appli

## ALIGNMENTS

### RESULT 1

US-09-862-027-27

; Sequence 27, Application US/09862027

; Patent No. US20020142428A1

; GENERAL INFORMATION:

; APPLICANT: Hodge, Martin R.

; TITLE OF INVENTION: No. US20020142428A1el Kinases and Uses Thereof

; FILE REFERENCE: 35800/234862

; CURRENT APPLICATION NUMBER: US/09/862, 027

; CURRENT FILING DATE: 2001-05-21

; PRIOR APPLICATION NUMBER: US 09/345,473

; PRIOR FILING DATE: 1999-06-30

; NUMBER OF SEQ ID NOS: 82

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 27

; LENGTH: 669

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

; US-09-862-027-27

Query Match	12.7%	Score 202.5;	DB 10;	Length 669;
Best Local Similarity	24.8%	Pred. No. 3.4e-11;		
Matches	80;	Conservative	63;	Mismatches 119;
				Indels 61;
				Gaps 14;
QY	6	FNFTKFDLDQKDLIFQGGDATSTNNVLQTLTKLDGNGPNVCASGVRLFSAPFHLWNSM--	63	
Db	28	FGFGNYLDNSGIAI---TNSKGLMKLT-----NSSEFSYGHVFNYSFVR-FKNSPNG	76	
QY	64	AVSFFETNLTQISTPHPYAAGDAFFLAPHDTVIPPNSNGKFLGLYSNVFNFRNSPTSEN	123	
Db	77	TVSSFTTTFVFAIVSNVNALDGHGLAFVISP-TKGLPYSSSSQYLGFLF-NLTNNGDPS--	132	
QY	124	QSGFDVNTDSRVAVFDTFEPNANIDP-NYRHIGIDVNSIKSKETARWEQ-----	173	
Db	133	-----NHIVAFDTFQNEFDDMDNNHVGIDINLSLSEKASTAGYEDDDGTGFKN	183	
QY	174	----NGKTATARIYNASAKKSTVT---TFYPGMEVWALSHDVLHAEPEWVRVGLSAS	226	
Db	184	IRLINQKPIQAWIEYDSSRRQLNVTIHPILPKPIPLSLTKDLSPLYFDSMYGVGFTSA	243	
QY	227	TGEKQKNTIISWTF--TSSLKNNEV-----KEPKEDMTIANVVRSYTWTINDVLSYISNKM	280	

Db 244 TGRLRSSHYILGWTFKNGTASNIDISRLPKLPDRSR--STSVKKILAILSLTSLAILV 301  
 QY 281 YDALNNHKKVRCSTCMFPMKKK 303  
 Db 302 FLTI-----SYMLFLKRRK 314

## RESULT 2

US-09-740-274-4  
 ; Sequence 4, Application US/09740274  
 ; Patent No. US20020031826A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Nichols, Scott E.  
 ; TITLE OF INVENTION: Glucan-containing Compositions and Paper  
 ; FILE REFERENCE: 0357CRD  
 ; CURRENT APPLICATION NUMBER: US/09/740,274  
 ; CURRENT FILING DATE: 2000-12-19  
 ; PRIOR APPLICATION NUMBER: 09/210,361  
 ; PRIOR FILING DATE: 1998-12-11  
 ; PRIOR APPLICATION NUMBER: 09/007,999  
 ; PRIOR FILING DATE: 1998-01-16  
 ; PRIOR APPLICATION NUMBER: 08/478,704  
 ; PRIOR FILING DATE: 1995-06-07  
 ; PRIOR APPLICATION NUMBER: 09/009,620  
 ; PRIOR FILING DATE: 1998-01-20  
 ; PRIOR APPLICATION NUMBER: 08/485,243  
 ; PRIOR FILING DATE: 1995-06-07  
 ; PRIOR APPLICATION NUMBER: 09/008,172  
 ; PRIOR FILING DATE: 1998-01-16  
 ; PRIOR APPLICATION NUMBER: 08/482,711  
 ; PRIOR FILING DATE: 1995-06-07  
 ; NUMBER OF SEQ ID NOS: 6  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 4  
 ; LENGTH: 1375  
 ; TYPE: PRT  
 ; ORGANISM: streptococcus mutans  
 ; US-09-740-274-4

Query Match 6.3%; Score 101; DB 10; Length 1375;  
 Best Local Similarity 23.6%; Pred. No. 0.51;  
 Matches 70; Conservative 42; Mismatches 104; Indels 80; Gaps 17;

QY 48 GRVLFSAFHLWNSMAVSSFTNTLIQISTP-----HPYYAAD----GFAFFLA--- 93  
 Db 385 GALLYS-----NNSKLSQANSYRIILNRTPTNQTGKDRPYTADRTTGGYEFLLANDV 438  
 QY 94 -PHDTVIPPN--SWGKELGLXSNVFNPSPTSENQSGFDVNTDS--RVVAVEFDTFPNANID 149  
 Db 439 DNSNPVQAEQLNLWHLFNMFGNIYANDP-----DANFDSIRVDAYD-----NVD 483  
 QY 150 PNYRHIGDVNSIKSKETARWQNGKTATARIYSNSA--SKXSTVTTFYPGMEVV----- 203  
 Db 484 ADLLQIAGDY-----LKAAGTKHKNDKAANDHLSILEAWSYNDTPYLHDDGDNMIMNDNR 538  
 QY 204 -----ALSHDVLHAELEPWVRVCLASSTGEEKKNTIISWSETSSLKNNVEKPEKED 256  
 Db 539 LRSLSLYLAKPLNQRSGNPLITNSLVNRTDQNAETAAVPSYFTRA--HDSVQD----- 593  
 QY 257 MYIANVRSYTWIN--DVLSY-----TSNKMYDALNNHKKVRCSTCMFLF 299  
 Db 594 -LIRNIIR--TEINPNVGVSYFTTEIKKAFETYNK--DLLATEKKYTHYNTALS 644

## RESULT 3

US-09-850-351A-32  
 ; Sequence 32, Application US/09850351A  
 ; Patent No. US2002010080A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Feitelson, Jerald S.  
 ; Schnepf, H. Ernest  
 ; Narva, Kenneth E.  
 ; Stockhoff, Brian A.

; Schmeits, James  
 ; Loewer, David  
 ; Dullum, Charles Joseph  
 ; Muller-Cohn, Judy  
 ; Stamp, Lisa  
 ; Morrill, George  
 ; TITLE OF INVENTION: No. US2002010080A1 Pesticidal Toxins and Nucleotide  
 ; Sequences Which Encode These Toxins  
 ; NUMBER OF SEQUENCES: 144  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
 ; STREET: 2421 N.W. 41st Street, Suite A-1  
 ; CITY: Gainesville  
 ; STATE: FL  
 ; COUNTRY: US  
 ; ZIP: 32606-6669  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/850,351A  
 ; FILING DATE: 07-May-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 09/073,898  
 ; FILING DATE: 06-MAY-1998  
 ; APPLICATION NUMBER: US 08/960,780  
 ; FILING DATE: 30-OCT-1997  
 ; APPLICATION NUMBER: US 60/029,848  
 ; FILING DATE: 30-OCT-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Sanders, Jay M.  
 ; REGISTRATION NUMBER: 39,355  
 ; REFERENCE/DOCKET NUMBER: MA-708CDI  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 352-375-8100  
 ; TELEFAX: 352-372-5800  
 ; INFORMATION FOR SEQ ID NO: 32:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 881 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; ORIGINAL SOURCE:  
 ; INDIVIDUAL ISOLATE: PSI77C8  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 32:  
 ; US-09-850-351A-32

Query Match 6.3%; Score 100; DB 10; Length 881;  
 Best Local Similarity 21.0%; Pred. No. 0.34; 133; Indels 80; Gaps 13;  
 Matches 66; Conservative 36; Mismatches 133; Indels 80; Gaps 13;  
 QY 9 TKFDLDOKDLIFQGDATSTNNVLQTLKLDGNNP-----VGASGRVLFSPAP- 55  
 Db 157 TKFNIDSK-----TPKELKLFKIDSNQPOQVQOODELRNPFNKKSEQELAKPS 206  
 QY 56 -FHLWNSMAVSSFTNTLIQISTPHYAAGDAF---FLAPHDTVIPPNWSGKFLGLY 111  
 Db 207 KINLFQKMKREIDEDTDTGDSIP-DLWEENGYTIQNIQIAVKWDDSLASKYTKFV--- 262  
 QY 112 SNVFNPSPTSENQSGFDVNTDSRVVAVEFD-----TFPNANI-----DPN 151  
 Db 263 -----SNPLESHVTGDPYTDYEKAARDLDSNAKETNPVLAAPFSPVNSMEKVLSPN 316  
 QY 152 YRHIGDVNSTKSKETARWQNGKTATARIYSNSASKSTVTTFYPGMEVVALSHDVL 211  
 Db 317 ENL-----SNSVESHSSTNWSYNTTEGASVEAGIGPKGISFGVSVNIQHSSETVA----- 365  
 QY 212 HAELEPWVRVCLASSTGEEKKNTIISWSETSSLKNNVEKPEKEDMYIANVRSYTWIND 271







TITLE OF INVENTION: NO. US2002010080A1el Pesticidal Toxins and Nucleotide Sequences Which Encode These Toxins

NUMBER OF SEQUENCES: 144

CORRESPONDENCE ADDRESS:

ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: US

ZIP: 32606-6669

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/850,351A

FILING DATE: 07-May-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/073,898

FILING DATE: 06-May-1998

APPLICATION NUMBER: US 08/960,780

FILING DATE: 30-OCT-1997

APPLICATION NUMBER: US 60/029,848

FILING DATE: 30-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Sanders, Jay M.

REGISTRATION NUMBER: 39,355

REFERENCE/DOCKET NUMBER: MA-708CD1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 352-375-8100

TELEFAX: 352-372-5800

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 425 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: 68F

SEQUENCE DESCRIPTION: SEQ ID NO: 27:

US-09-850-351A-27

Query Match 5.6% Score 89; DB 10; Length 425;

Best Local Similarity 20.6%; Pred. No. 1.4; Indels 90; Gaps 12;

Matches 59; Conservative 31; Mismatches 106; Indels 90; Gaps 12;

QY 20 FQGDATSTNNVQLTK-----LDSGGNPGVGSVGRVLFSAFPHLWENSMVSSPETNLT 73

Db 152 FLTKATKTNLTQVKSTRDEDTDGD-----SIP-DIWE-----ENGYT 191

QY 74 IQISTPHY---YAADGFAFLA-PHDTVIPNPGWKGFLGLYSNVFRNSPTSENSFGDV 129

Db 192 IQNKIAVKWDBSLASKGYTKFVSNPLDT-----HTVGDP 225

QY 130 NTDSRVVAFED-----TFPNAIDPNYRHIGID---VNSIKSKETARWQ 173

Db 226 YTDEKAAARDLDLSNAKETFNPLVAAFPSVNVSMKEVILSPDENLSNIESHSSTNWSYT 285

QY 174 NGKTATARISYNSASKSTVTTFYPGMEVVALSHDVLHAEPEWVRVGLSASTGEKQK 233

Db 286 NTEGASIEAGGALGLFGVSANTQHSETVGY-----EW-----GTSTGNTSQF 329

QY 234 NTIISWFTSLKKNVEKPEDMYIANVRSYTWINDVLSYISNK 279

Db 330 NTASAGYLNANVRNVGTGA--IYDVKPTTSFVLNKKDTIATITAK 373

RESULT 10

US-10-192-584-6

; Sequence 6, Application US/10192584

; Publication No. US20030027987A1

; GENERAL INFORMATION:

; APPLICANT: TOKUNAGA, Eiji

; SAKAGUCHI, Masashi

; MATSUO, Kazuo

; HAMADA, Fukusaburo

; TOKIYOSHI, Sachio

; TITLE OF INVENTION: NOVEL POLYPEPTIDE FROM HAEMOPHILUS

; PARAGALLINARUM AND PROCESS FOR PREPARING THE SAME

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BROWDY AND NEIMARK

; STREET: 624 Ninth Street, N.W., Suite 300

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20001

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/192,584

; FILING DATE: 11-Jul-2002

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/077,098

; FILING DATE: 19-May-1998

; APPLICATION NUMBER: PCT/JP97/03222

; FILING DATE: 12-SEP-1997

; APPLICATION NUMBER: JP 27,148/1996

; FILING DATE: 19-SEP-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: KORNBAU, Anne M.

; REGISTRATION NUMBER: 25,618

; REFERENCE/DOCKET NUMBER: TOKUNAGA-1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-628-5197

; TELEFAX: 202-737-3528

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2042 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-10-192-584-6

Query Match 5.5% Score 88; DB 9; Length 2042;

Best Local Similarity 22.4%; Pred. No. 16;

Matches 62; Conservative 42; Mismatches 107; Indels 66; Gaps 13;

QY 22 GDATS---TNVLQLTKLDSGGNPGVGSVGRVLFSAFPHLWENSMVSSFETNLTQIST 78

Db 1209 GDANNKQLNNTLQQTLEATG--ITSVSGSTNY-AGFSLGADSVTFSGGAG-TVKLSG 1264

QY 79 PHPYAAADGFA-----FFLAPHDTVIPNPGWKGFLGLYSNVFRNSPTSENSQSGD 128

Db 1265 VSDATADTAATLKQVKEYRTTLVGDNDITAAADRSGGTSNGITYNLSLNGKTVS----- 1318

QY 129 VNTDSRVVA--VEFDIFPNANIDPNYRHIGID---VNSIKSKETARWQNGKTATARIS 183

Db 1319 -ATEEKVSVGKTVYEAIRNA-ITGNIFTGLDDTTLNKNIN-----PADQDLS 1364

QY 184 YNSASKKSTVTTFYPGMEVVALSHDVLHAEPEWVRVGLSASTGEKQKQKTIISWSETS 243

Db 1365 NLSESGKNAITGL--VDVVVKKINS-----PITVEPSTDSNKKKTFVGVDFD 1410

QY 244 SLKNNVEKPEKEDMYIANVRSYTWINDVLSYISNM 280

Db 1411 TITEGDATDDK-----LTTSKSVESYVTNKL 1437

## RESULT 11

US-09-815-242-5763  
; Sequence 5763, Application US/09815242  
; Patent No. US20020061569A1

## ; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.

; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO: 5763

; LENGTH: 234

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-815-242-5763

Query Match 5.5%; Score 87.5; DB 10; Length 234;

Best Local Similarity 23.7%; Pred No. 0.84; 79; Indels 59; Gaps 11;

Matches 51; Conservative 26; Mismatches 79; Indels 59; Gaps 11;

QY 9 TKFDLDQKDLIFQGDATSTNNVLQTLKDSGPNVGASVGRVLFSAFFHLWENSMVSSFG 68

Db 4 TNLNVRDLDM--TTFYKNILGLSVKSSDDNTTVLSVGTGGHTLTLLHLEDGRTSPR 61

QY 69 ETNL-TIQISTPHPYAAGDGAFFLAHPDHTVIPPNSWKGKFLGLYSNVFNSPTSENQSGF 127

Db 62 EAGLFHIAFLPTTDLAN-FLYFVAQ-----KNMGIGAG---DHLVSEALYFN 106

QY 128 DVNTDSRVVAVEFDTPFNANIDPNYRHIGIDVNSIKSETARWQNGKTATARISYNSA 187

Db 107 D-----PEGN-----GIEV---YDRPSSSWWQNGKVKMDTLEVD- 140

QY 188 SKKSTVTF-----YPMGEVVAL-----SHDVD 210

Db 141 ---QTLTTHRTDEGWQGMKPAKGMIHGLHLKTHDLD 172

## RESULT 12

US-09-815-242-12678

; Sequence 12678, Application US/09815242

; Patent No. US20020061569A1

## ; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO: 12678

; LENGTH: 268

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-815-242-12678

Query Match 5.5%; Score 87.5; DB 10; Length 268;

Best Local Similarity 23.7%; Pred No. 1;

Matches 51; Conservative 26; Mismatches 79; Indels 59; Gaps 11;

QY 9 TKFDLDQKDLIFQGDATSTNNVLQTLKDSGPNVGASVGRVLFSAFFHLWENSMVSSFG 68

Db 12 TNLNVRDLDM--TTFYKNILGLSVKSSDDNTTVLSVGTGGHTLTLLHLEDGRTSPR 69

QY 69 ETNL-TIQISTPHPYAAGDGAFFLAHPDHTVIPPNSWKGKFLGLYSNVFNSPTSENQSGF 127

Db 70 EAGLFHIAFLPTTDLAN-FLYFVAQ-----KNMGIGAG---DHLVSEALYFN 114

QY 128 DVNTDSRVVAVEFDTPFNANIDPNYRHIGIDVNSIKSETARWQNGKTATARISYNSA 187

Db 115 D-----PEGN-----GIEV---YDRPSSSWWQNGKVKMDTLEVD- 148

QY 188 SKKSTVTF-----YPMGEVVAL-----SHDVD 210

Db 149 ---QTLTTHRTDEGWQGMKPAKGMIHGLHLKTHDLD 180

## RESULT 13

US-09-853-533A-8

; Sequence 8, Application US/09853533A

; Patent No. US20020103362A1

## ; GENERAL INFORMATION:

; APPLICANT: Monsanto Company

; APPLICANT: Isaac, Barbara

; APPLICANT: Krieger, Elysia

; APPLICANT: Mettus, Anne-Marie

; APPLICANT: Moshiri, Farhad

; APPLICANT: Sivasubramaniam, Sakuntala

; TITLE OF INVENTION: POLYPEPTIDE COMPOSITIONS TOXIC TO ANTHONOMUS INSECTS, AND METH

; FILE REFERENCE: 38-21(51932)B

; CURRENT APPLICATION NUMBER: US/09/853,533A

; CURRENT FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: 60/204,367

; PRIOR FILING DATE: 2000-05-15

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO: 8

; LENGTH: 632

; TYPE: PRT

; ORGANISM: Bacillus thuringiensis





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:40:30 : Search time 37.0712 Seconds  
(without alignments)  
1089.120 Million cell updates/sec

Title: US-09-476-485A-6  
Perfect score: 1599  
Sequence: 1 AQLSLFNFTRFDLQDKLIF.....LNNHKKVRCSTCMFMKKK 303

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1599	100.0	303	22	AAG62898
2	804	50.3	264	20	AAW87973
3	804	50.3	264	22	AAG62890
4	804	50.3	286	22	AAG62894
5	787.5	49.2	234	22	AAG62901
6	658	41.2	290	21	AAV58736
7	653	40.8	290	21	AAV58737
8	564	35.3	237	16	AAW74765
9	493.5	30.9	285	20	AAV06811
10	489.5	30.6	285	20	AAV06814

11	485.5	30.4	285	20	AAV06812	R. pseudoacacia le
12	483.5	30.2	285	20	AAV06815	Amino acid sequenc
13	471	29.5	275	21	AAV58738	Pea lectin, 2ltn.
14	416	26.0	228	16	AAW74766	Pea lectin, 2ltn.
15	383	24.0	262	20	AAW27265	B. purpurea native
16	378	23.6	242	15	AAW45911	Bauhinia purpurea l
17	373	23.3	242	15	AAW45912	Bauhinia purpurea l
18	350	21.9	236	16	AAW64807	PNA lectin subunit
19	345.5	21.6	239	16	AAW74767	Pea lectin, 1ltn.
20	308.5	19.3	274	20	AAV06813	R. pseudoacacia le
21	287	17.9	265	10	AAW91967	Sequence of arceli
22	250.5	15.7	282	21	AAW25452	Pinus radiata cell
23	245	15.3	652	23	ABW93374	Herbicidally activ
24	243	15.2	681	23	ABW93373	Herbicidally activ
25	237.5	14.9	244	10	AAW93640	Sequence of a Phas
26	221.5	13.9	632	21	ABW25109	Pinus radiata cell
27	209.5	13.1	715	23	ABW92649	Herbicidally activ
28	203.5	12.7	283	23	ABW91344	Herbicidally activ
29	203	12.7	711	23	ABW93337	Herbicidally activ
30	202.5	12.7	669	23	ABW93342	Herbicidally activ
31	198	12.4	649	23	ABW93334	Herbicidally activ
32	197.5	12.4	675	23	ABW92029	Herbicidally activ
33	197.5	12.4	677	23	ABW92654	Herbicidally activ
34	195.5	12.2	674	23	ABW93878	Herbicidally activ
35	193.5	12.1	691	23	ABW93681	Herbicidally activ
36	188.5	11.8	272	23	ABW91343	Herbicidally activ
37	186	11.6	242	23	ABW91342	Herbicidally activ
38	185	11.6	604	23	ABW92522	Herbicidally activ
39	179.5	11.2	666	23	ABW91530	Herbicidally activ
40	177	11.1	623	23	ABW91919	Herbicidally activ
41	177	11.1	682	23	ABW92521	Herbicidally activ
42	176.5	11.0	693	23	ABW92247	Herbicidally activ
43	169.5	10.6	685	23	ABW93319	Herbicidally activ
44	167	10.4	657	23	ABW93898	Herbicidally activ
45	162.5	10.2	627	23	ABW91918	Herbicidally activ

## ALIGNMENTS

RESULT 1  
AAG62898  
ID AAG62898 standard; Protein; 303 AA.  
AC AAG62898;  
XX  
XX  
XX  
DT 17-SEP-2001 (first entry)  
XX  
XX  
DE Amino acid sequence of a french bean FRIL polypeptide.  
XX  
XX  
KW FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;  
KW progenitor cell preservation factor; radiotherapy; chemotherapy;  
KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
KW severe combined immunodeficiency; aplastic anemia; tissue repair.  
OS Phaseolus vulgaris.  
XX  
XX  
PN WO200149851-A1.  
XX  
PD 12-JUL-2001.  
XX  
PF 30-DEC-1999; 99WO-US31307.  
XX  
PR 30-DEC-1999; 99WO-US31307.  
XX  
PA (PHYL-) PHYLOGIX LLC.  
XX  
PI Colucci MG, Chrispeels MJ, Moore JG;  
XX  
DR WPI; 2001-441882/47.  
DR N-PSDB; AAH42306.  
XX  
PT Legume Progenitor cell preservation factors for in vivo or ex vivo

PT preservation of hematopoietic progenitor cells and as therapeutics for  
PT alleviating/reducing progenitor cell-depleting activity of cancer  
XX therapeutics  
XX  
XX Example 5; Page 81; 173pp; English.  
XX  
CC The present sequence represents a FRIL (FLK2/FLT3 tyrosine kinase  
CC receptor-interacting lectin) polypeptide. The specification describes a  
CC composition of one or more members of FRIL family of progenitor cell  
CC preservation factors. The composition is useful for alleviating or  
CC reducing the hematopoietic progenitor cell-depleting activity of  
CC a therapeutic treatment, including radiotherapeutic and/or  
CC chemotherapeutic treatments. Administration of FRIL compositions to a  
CC patient prior to treatment of the patient with a therapeutic treatment  
CC having a hematopoietic progenitor cell-depleting activity alleviates or  
CC reduces the hematopoietic progenitor cell-depleting activity of the  
CC therapeutic treatment in the patient. FRIL family members are useful for  
CC isolating population of progenitor cells, hemangioblasts, and mesenchymal  
CC stem cells. The composition is administered to reduce progenitor cell  
CC depleting effects of chemotherapeutics, so that the patient can receive  
CC a higher dose of the chemotherapeutic and preferably recover from cancer.  
CC It is also administered to patients having, or predisposed to developing  
CC a condition where the patients hematopoietic progenitor cells are  
CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
CC The isolated mesenchymal cells are useful for tissue repair.

XX Sequence 303 AA;

Query Match 100.0%; Score 1599; DB 22; Length 303;  
Best Local Similarity 100.0%; Pred. No. 2.3e-141; Indels 0; Gaps 0;  
Matches 303; Conservative 0; Mismatches 0;

Qy 1 AQSLSFNFTKFDLDKDLIFOGDATSTNNVLQTLKLDGSGNPVGASVGRVLSAPFHLWE 60  
Db 1 AQSLSFNFTKFDLDKDLIFOGDATSTNNVLQTLKLDGSGNPVGASVGRVLSAPFHLWE 60  
Qy 61 NSMAVSFETNLTIQISTPHYPYAADGFAFFLPHDVTIPPNSWKGFLGLYSNVFRNSPT 120  
Db 61 NSMAVSFETNLTIQISTPHYPYAADGFAFFLPHDVTIPPNSWKGFLGLYSNVFRNSPT 120  
Qy 121 SENQSGDVNTDSRVVAVVEFDFFPNANIDPNYRHIGIDVNSIKSKETARWQNGKTATA 180  
Db 121 SENQSGDVNTDSRVVAVVEFDFFPNANIDPNYRHIGIDVNSIKSKETARWQNGKTATA 180  
Qy 181 RISYNSASKKSTVTTFYPCGMEVVALSHVDLHAELPEWVRVGLSASTGEEKOKNTIISWS 240  
Db 181 RISYNSASKKSTVTTFYPCGMEVVALSHVDLHAELPEWVRVGLSASTGEEKOKNTIISWS 240  
Qy 241 FTSSLKNNVEKPEKEDMYIANVVRSTWINDVLSYISNKKMYDALNNHKKYVRCSTCMLFM 300  
Db 241 FTSSLKNNVEKPEKEDMYIANVVRSTWINDVLSYISNKKMYDALNNHKKYVRCSTCMLFM 300

Qy 301 KKK 303  
Db 301 KKK 303

RESULT 2  
ID AAW87973  
XX AAW87973 standard; Protein; 264 AA.  
AC AAW87973;  
XX  
XX  
DT 13-APR-1999 (first entry)  
DE A lectin derived progenitor cell preservation factor.

KW Lectin derived progenitor cell preservation factor; progenitor cell;  
KW hematopoietic cell; cultured cell preservation; anticancer therapy;  
KW myeloablative therapy; sickle-cell anaemia; ablative therapy protection;  
KW FLK2/FLT3 receptor.  
XX Dolichos lab lab.

XX  
PN WO9859038-A1.  
XX  
PD 30-DEC-1998.  
XX  
PF 23-JUN-1998; 98WO-US13046.  
XX  
PR 24-JUN-1997; 97US-0881189.  
XX  
PA (IMCL-) IMCLONE SYSTEMS INC.  
PA (REGC) UNIV CALIFORNIA.  
XX  
PI Chrispeels MJ, Colucci MG, Moore JG;  
XX WPI; 1999-081274/07.  
DR N-PSDB; AAX03593.  
XX  
XX New nucleic acid encoding plant lectin that preserves progenitor  
PT cells - particularly haematopoietic progenitors, useful for bone  
PT marrow reconstitution after ablative therapy, and to increase DNA  
PT transfer in gene therapy  
XX  
PS Claim 1; Page 30-31; 72pp; English.

CC The present sequence represents a lectin derived progenitor cell  
CC preservation factor. The protein is used to preserve unipotent,  
CC pluripotent or totipotent progenitor cells, especially haematopoietic  
CC cells, and also progenitors from nerve, muscle, skin, gut, bone,  
CC kidney, liver, pancreas or thymus. Specific applications are  
CC preservation of cultured cells intended for administration after  
CC (anticancer) myeloablative therapy (bone marrow or whole-body irradiation  
CC or chemotherapy) to reconstitute the haematopoietic system; enrichment  
CC of progenitor cells (e.g. during ex vivo purging of malignant cells);  
CC treatment of tissues containing haematopoietic progenitors for subsequent  
CC transplant to improve haematopoietic competence; improving transfer of  
CC exogenous DNA to progenitor cells (in gene therapy of various  
CC haematological disorders, e.g. sickle-cell anaemia); and protection  
CC against ablative therapy (to eliminate proliferating cells specifically),  
CC followed by re-establishment of differentiation and proliferation of  
CC preserved progenitors. The protein, when linked to magnetic beads, may  
CC also be used to isolate cells that express the FLK2/FLT3 receptor.

XX Sequence 264 AA;

Query Match 50.3%; Score 804; DB 20; Length 264;  
Best Local Similarity 60.9%; Pred. No. 6.3e-67;  
Matches 162; Conservative 39; Mismatches 59; Indels 6; Gaps 4;

Qy 1 AQSLSFNFTKFDLDKDLIFOGDATSTNNVLQTLKLDGSGNPVGASVGRVLSAPFHLWE 60  
Db 1 AQSLSFNFTKFDLDKDLIFOGDATSTNNVLQTLKLDGSGNPVGASVGRVLSAPLRLWE 60  
Qy 61 NSMAVSFETNLTIQISTPHYPYAADGFAFFLPHDVTIPPNSWKGFLGLY--SNVFRNS 118  
Db 61 DSAVLTSFDITINFEISTPTYSRIADGLAFFIAPDPSVI--SYHGFGFLGFPFNANLNS 118  
Qy 119 PTSENQSGDVNTDSRVVAVVEFDFFPNANI-DPNYRHIGIDVNSIKSKETARWQNGKT 177  
Db 119 STSENQTTKA-ASSNVVAVVEFDYLNPDYGNPVIHIGIDVNSIRSKVTAKDWQNGKI 177  
Qy 178 ATARISYNSASKKSTVTTFYPCGMEVVALSHVDLHAELPEWVRVGLSASTGEEKOKNTII 237  
Db 178 ATAHISYNSVSKRLSVTSYAGSKPATLSYDIELHTLVLPWVRVGLSASTGQDKERTNVH 237  
Qy 238 SWSFTSSLKNNVEKPEKEDMYIANVV 263  
Db 238 SWSFTSSLTWNVAKKENENKVIITRGV 263

RESULT 3  
ID AAG62890  
XX AAG62890 standard; Protein; 264 AA.  
XX



AC AAG62890;  
 XX 17-SEP-2001 (first entry)  
 XX Amino acid sequence of a hyacinth bean FRIL polypeptide.  
 DE  
 XX FRIL; FIK2/Flt3 tyrosine kinase receptor-interacting lectin;  
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;  
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
 KW severe combined immunodeficiency; aplastic anemia; tissue repair.  
 XX  
 OS Dolichos lab lab.  
 XX WO200149851-A1.  
 PN 12-JUL-2001.  
 XX 30-DEC-1999; 99WO-US31307.  
 XX 30-DEC-1999; 99WO-US31307.  
 PR (PHYL-) PHYLOGIX LLC.  
 XX Colucci MG, Chrispeels MJ, Moore JG;  
 XX WPI; 2001-441882/47.  
 DR N-PSDB; AAH42287.  
 XX Legume progenitor cell preservation factors for in vivo or ex vivo  
 PT preservation of hematopoietic progenitor cells and as therapeutics for  
 PT alleviating/reducing progenitor cell-depleting activity of cancer  
 PT therapeutics  
 XX  
 PS Example 1; Page 54-55; 173pp; English.  
 XX  
 CC The present sequence represents a FRIL (Flk2/Flt3 tyrosine kinase  
 CC receptor-interacting lectin) polypeptide. The specification describes a  
 CC composition of one or more members of FRIL family of progenitor cell  
 CC preserving factors. The composition is useful for alleviating or  
 CC reducing the hematopoietic progenitor cell-depleting activity of  
 CC a therapeutic treatment, including radiotherapeutic and/or  
 CC chemotherapeutic treatments. Administration of FRIL compositions to a  
 CC patient prior to treatment of the patient with a therapeutic treatment  
 CC having a hematopoietic progenitor cell-depleting activity alleviates or  
 CC reduces the hematopoietic progenitor cell-depleting activity of the  
 CC therapeutic treatment in the patient. FRIL family members are useful for  
 CC isolating population of progenitor cells, hemangioblasts, and mesenchymal  
 CC stem cells. The composition is administered to reduce progenitor cell  
 CC depleting effects of chemotherapeutics, so that the patient can receive  
 CC a higher dose of the chemotherapeutic and preferably recover from cancer.  
 CC It is also administered to patients having, or predisposed to developing  
 CC a condition where the patients hematopoietic progenitor cells are  
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
 CC The isolated mesenchymal cells are useful for tissue repair.  
 XX  
 SQ Sequence 264 AA;

Query Match 50.3%; Score 804; DB 22; Length 264;  
 Best Local Similarity 60.9%; Pred. No. 6.3e-67;  
 Matches 162; Conservative 39; Mismatches 59; Indels 6; Gaps 4;

QY 1 AQSLSFNFTKDLDDKLLIFQGDATSTNNVQLTKLDSGGNPGVSGVRFSLFAPHLWE 60  
 DB 1 AQSLSFSFTKDFDNPQBDLIFQGHATSTNNVQLTKLDSAGNPVSSSAGRVLYSAPLRWE 60  
 QY 61 NSMAVSSFTNLTIQISTPHPPYAAADGFAFFLPHDPTVIPPNSGKFLGLY--SNVFN 118  
 DB 61 DSAVLTSFDTIINFELSTYTSRIAGLAFIAPPDSVI--SYHGFLGLFPNPNLUNNS 118  
 QY 119 PRSENGSGDVNTDSRVAVEFTFPNANI-DPNYRHIGIDVNSIKSKETARWEQNGKT 177  
 DB 119 STSENQTTTKA-ASSNVVAVEFTYLPDYGDPNYIHIGIDVNSIRSKVTAKWDQNGKI 177

QY 178 ATARISYNSASKKSTVTTFYPCMEVVALSHDVLHAEPEWVRVGLSASTGEQKNTII 237  
 DB 178 ATAHISYNSVSKRLSVTSYAGSKPATLSYDIELHTVLPWVRVGLSASTGQDKERTVH 237  
 QY 238 SWSFTSSLKKNVEKPEKEDMYIANVV 263  
 DB 238 SWSFTSSLTNTVAKKENENKYYIRGV 263  
 RESULT 4  
 AAG62894  
 ID AAG62894 standard; Protein; 286 AA.  
 XX AC AAG62894;  
 XX DT 17-SEP-2001 (first entry)  
 XX DE Alpha-amylase inhibitor signal peptide and FRIL fusion.  
 XX KW FRIL; FIK2/Flt3 tyrosine kinase receptor-interacting lectin;  
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;  
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
 KW severe combined immunodeficiency; aplastic anemia; tissue repair;  
 KW alpha-amylase inhibitor gene.  
 XX OS Synthetic.  
 OS Unidentified.  
 OS Dolichos lab lab.  
 PN WO200149851-A1.  
 XX 12-JUL-2001.  
 XX 30-DEC-1999; 99WO-US31307.  
 XX 30-DEC-1999; 99WO-US31307.  
 PA (PHYL-) PHYLOGIX LLC.  
 PT Colucci MG, Chrispeels MJ, Moore JG;  
 XX WPI; 2001-441882/47.  
 DR N-PSDB; AAH42295.  
 XX Legume progenitor cell preservation factors for in vivo or ex vivo  
 PT preservation of hematopoietic progenitor cells and as therapeutics for  
 PT alleviating/reducing progenitor cell-depleting activity of cancer  
 PT therapeutics  
 XX  
 XX Example 1; Page 59; 173pp; English.  
 CC The present sequence represents fusion protein of alpha-amylase inhibitor  
 CC signal peptide and FRIL (Flk2/Flt3 tyrosine kinase receptor-interacting  
 CC lectin). The specification describes a composition of one or more members  
 CC of FRIL family of progenitor cell preservation factors. The composition  
 CC is useful for alleviating or reducing the hematopoietic progenitor  
 CC cell-depleting activity of a therapeutic treatment, including  
 CC radiotherapeutic and/or chemotherapeutic treatments. Administration of  
 CC FRIL compositions to a patient prior to treatment of the patient with  
 CC a therapeutic treatment having a hematopoietic progenitor cell-depleting  
 CC activity alleviates or reduces the hematopoietic progenitor  
 CC cell-depleting activity of the therapeutic treatment in the patient.  
 CC FRIL family members are useful for isolating population of progenitor  
 CC cells, hemangioblasts, and mesenchymal stem cells. The composition is  
 CC administered to reduce progenitor cell depleting effects of  
 CC chemotherapeutics, so that the patient can receive a higher dose of the  
 CC chemotherapeutic and preferably recover from cancer. It is also  
 CC administered to patients having, or predisposed to developing a  
 CC condition where the patients hematopoietic progenitor cells are  
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
 CC The isolated mesenchymal cells are useful for tissue repair.  
 XX  
 SQ Sequence 286 AA;



The present sequence is that of the concanavalin A (Con A) lectin of jack bean. The present invention relates to: a transformed *Brassica* plant that is resistant to certain insect pests; an expression cassette containing DNA that codes for at least 1 lectin selected from Con A, modified Con A (see AY58737) and pea lectin (see AY58738); transgenic plant cells containing at least 1 copy of the DNA; a new lectin derived from the jack bean Con A gene; a method of imparting resistance to insects selected from blossom beetles (pollen beetles) of the genus *Meligethes*, flea beetles of the genus *Phyllotreta*, and root flies of the genus *Delia*; and a method for protecting a plant against infestation by insects of these genera.

[illegible]

RESULT 7	
AA58737	
ID	AA58737 standard; Protein; 290 AA.
XX	
XX	AA58737;
XX	
XX	25-APR-2000 (first entry)
DT	
XX	
XX	Jack bean concanavalin A N152S mutant.
DE	
XX	
XX	Concanavalin A; Con A; lectin; jack bean; insecticide;
KW	transgenic plant; Brassica; insect resistance; mutain.
XX	
XX	Canavalia ensiformis.
OS	

	Key	Location/Qualifiers
..		
FH		
FT	misc_difference	152
FT		/note= "replaces wild-type Asn"

WO200001223-A1.  
13-JAN-2000.

XX		
PF	02-JUL-1999;	99WO-SE01209.
XX		
PR	03-JUL-1998;	98SE-0002425.

XX (PLAN-) PLANT SCI SVERIGE AB.

PI Ahman I, Melander M, Vamling K:

XX  
DR WPI; 2000-160693/14.

DR N-PSDB; AAZ58018.

XX PT PT XX PS XX CC CC CC CC CC CC CC CC CC CC CC CC CC CC CC XX SQ

Query Match	40.8%	Score 653;	DB 21;	Length 290;
Best Local Similarity	51.9%	Pred. No. 1e-52;		
Matches 139; Conservative	43;	Mismatches 66;	Indels 20;	Gaps 9

QY	3	SLSEFTKFDLQDKDLIFQGDATS--TNNVLQLTKLDSGGNPVGASVGRVLFSAPFHLWEN	61
Db	36	ALHFMNQFSKDKDLILQDGTATGTGDNLELFRVSSNGSPQSSVGRALFYAPVHIWES	95
QY	62	SMVSSSFETNLTIOISTP--HPYVADGGAFFLAPHDVTIPPNSCKEFLGLY--SNVFRN	117
Db	96	SAVVASEATFTFLIKSPDHP---ADGIAFFISNIDSSIPSGTGRLLGLFPDANVRS	152
QY	118	SPTSE--NQSGDVNTDSRVVAEFDTFPPNANI--DPNRYRHIGIDVNSIKSKETARWQMNG	175
Db	153	STATIDFNAAY--NADT-IVAVELDTYPNTDIDGDPSPYHIGIDIKSVRSKKTAKNMNG	208
QY	176	KTATARISYNSAKSKSTVTTFYPGMEVVALSHDVLHAELPFWVRVGLSASTCEEKQKNT	235
Db	209	KVCTAHILYNSVDKRLSAYVSYPNADSATVSYDVLNDLVLPFWVRVGLSASTGLYKETNT	268
QY	236	IISWSTTSLSLKNNEVKPEKEDMYIANVV	263
Db	269	ILSWSTFSKLSKNEIPD-----IATVV	290

RESULT 8  
AAR74765  
ID AAR74765 standard: protein: 237 AA.

AA  
AC AAR74765;

DT 13-OCT-1995 (first entry)

DE Lequme concanavalin A.

xx Pentraxin; legume; lectin; structural homology; serum amyloid P; SAP;  
 xx amyloid; fibril; human; Limulus; C-reactive protein; CRP; pea;  
 kw female hamster protein; 3-D structure; jellyroll topology;  
 kw plcornavirus; coat protein; pentameric structure; concanavaline A.  
 kw

OS Legume sp..

PN WO9505394-A.

AA  
PD  
23-FEB-1995.

XX  
PF 17-AUG-1994; 94WO-GB01802.

XX  
PR 17-AUG-1993; 93GB-0017120.

XX PA (BIRK-) BIRKBECK COLLEGE.

(ROYA-) ROYAL POSTGRAD MED SCHOOL.

PA

Blundell TL, Pepys MB;

XX

WPI; 1995-098720/13.

XX

New cpds. inhibiting binding of serum amyloid P to amyloid fibrils - produced by computer assisted molecular design, useful for preventing, treating or diagnosing amyloidosis, e.g. Alzheimer's disease

XX

Claim 7; Fig 4; 72pp; English.

XX

The sequences given in AAR74763-70 represent various pentraxins and legume lectins. All these proteins show structural homology and may be used in the molecular design of a molecule for the inhibition of serum amyloid P (SAP) binding to amyloid fibrils. The similarities in the amino acid sequences of SAP, human and Limulus C-reactive protein (CRP) and female hamster protein suggests that they may have similar 3-D structures. The jellyroll topology of the pentraxins is reminiscent of the picornavirus coat proteins which also have pentameric structures. However, pentaxins resemble more closely legume lectins such as concanavaline A and pea lectin. Alignment of sequences on the basis of topologically equivalent features of the three dimensional structures shows that helices occupy different positions in the pentraxins and legume lectins and that the amino acid sequences of the two families have identities of only approx. 11%. The two main helices in SAP occur before and after strand L, whereas the helices in the legume lectins occur at the C-terminus of strand J. There is a long insertion between the end of the helix after strand D and the beginning of strand E in the lectins relative to pentraxins. Strands G, H and I together with the type IV beta-hairpin between H and I are identical in both SAP and pea lectin. The so-called pentraxin octapeptide signature sequence, HXCXS/TWXS, is in this region so it is not conserved in the legume lectins.

XX SQ Sequence 237 AA;

Query Match 35.3%; Score 564; DB 16; Length 237;

Best Local Similarity 49.8%; Pred. No. 1.7e-44;

Matches 115; Conservative 34; Mismatches 60; Indels 22; Gaps 5;

QY 25 TSTNNVLQTLKLDGSGNPGVSGVRLFSAPHLWNSMAVSSFFETNLTIQISTP--HPY 82

Db 2 TGTGDNLELTRVSSNGSGVSGVRLFSAPHLWNSMAVSSFFETNLTIQISTP--HPY 82

QY 83 YAADGFAFFLAPHDVTIPPNSMGKFLGYSNVFRNSPTSENQSGDVNTDSRVVAVFEFT 142

Db 61 --ADGIAFFISNIDSSIPSGSTGRILGL-----FPDANADT-IVAVELDT 102

QY 143 FPNANI-DPNYRHIGIDVNSIKSKETARWONGKTATARIYSNSASKSTVTTFYPGME 201

Db 103 YPNTDIDGPSYPHIGIDIKSVRSKKTAKNMQDGKGTAAHIIYNSVDKRLSVAVSPNAD 162

QY 202 VVALSHVDLHAELEPWVVGLSASGEKOKNTIISWFSSTSLKKNVKE 252

Db 163 ATSVSYVDLNDLVPWVRVGLSASTGLYKETNTIISWFSSTSLKKNSTHQ 213

RESULT 9

AAY06811

ID AAY06811 standard; protein; 285 AA.

XX AC AAY06811;

XX AC AAY06811;

DT 23-JUN-1999 (first entry)

DE R. pseudoacacia lectin RPBAl polypeptide subunit a.

XX R. pseudoacacia; lectin; medicine; mucosal cell proliferation;

XX R. pseudoacacia; lectin; medicine; mucosal cell proliferation;

XX cancer; metabolic disorder; mucositis; cytotoxic.

XX cancer; metabolic disorder; mucositis; cytotoxic.

OS Robinia pseudoacacia.

XX

WO9911278-A1.

XX

11-MAR-1999.

XX

28-AUG-1998; 98WO-GB02612.

XX

29-AUG-1997; 97GB-0018413.

XX

(ALIZ-) ALIZYME THERAPEUTICS LTD.

XX

Bardocz ZM, Fish NW, Palmer RMJ, Pusztai AJ;

XX

WPI; 1999-228935/19.

XX

Use of Robinia pseudoacacia lectin in medicine

XX

Disclosure; Fig 1; 37pp; English.

XX

The invention relates to Robinia pseudoacacia lectins which can be used in medicine. The lectin is used for the control of mucosal cell proliferation, for the reduction and/or treatment of damage caused by cell damaging agents, especially in the treatment of cancer, and/or for the reduction and/or treatment of metabolic disorders. It is especially useful for the treatment of mucositis in mammalian cells and/or tissues, particularly human cells and/or tissues (especially mucous cells including mucous membrane). Use of the lectin is effective and does not damage the gut as is the case with cytotoxic drugs and radiation. The present sequence represents a polypeptide subunit of R. pseudoacacia lectin RPBAl.

XX SQ Sequence 285 AA;

Query Match 30.9%; Score 493.5; DB 20; Length 285;

Best Local Similarity 44.4%; Pred. No. 9.1e-38;

Matches 115; Conservative 38; Mismatches 83; Indels 23; Gaps 8;

QY 3 SLSNFTKFDLDQKDLIFQGDATSTN-NVLQTLKLDGSGNPGVSGVRLFSAPHLWEN 61

Db 34 SLSFSFPKFAFNQPLIFQGDALVTSTGLQLTNV-VNGVSPGKSLGALYAAPQINDS 92

QY 62 SMA-VSSFETNLTIQISTPFPYAADGFAFFLAPHDVTIPPNSMGKFLGYSNVFRNSPT 120

Db 93 TTGNVASEVTSFSLIQAPNTTTADGLAFPLAPVD--PQLDVGMLGIFKD----- 143

QY 121 SENOSFGDVNTDSRVVAVFEFTFPNANIDPNYRHIGIDVNSIKSKETARWONGKTATA 180

Db 144 -----GVFNKSQIIVAVEFDTFSNIHFDKGRHMGINNVSVISIKTPVNNWNGEVANV 197

QY 181 RISYNSASKSTVTTFYPGMEVVALSHD-VDLHAELEPWVVRVGLSASTGEK--QKNTI 236

Db 198 FISYEASTKSLTASLVPSLETSTFIVHAIVDVQVLPWVRVGFSGATTGIDGKYVQINDV 257

QY 237 ISWSFTSSLK-NNEVKEPK 254

Db 258 LSWSFENLPGNSVASVK 276

RESULT 10

AAY06814

ID AAY06814 standard; protein; 285 AA.

XX AC AAY06814;

XX AC AAY06814;

DT 23-JUN-1999 (first entry)

DE Amino acid sequence of lectin RPBAl from R. pseudoacacia seed.

XX Robinia pseudoacacia; lectin; medicine; mucosal cell proliferation;

XX Robinia pseudoacacia; lectin; medicine; mucosal cell proliferation;

XX cancer; metabolic disorder; mucositis; cytotoxic.

XX cancer; metabolic disorder; mucositis; cytotoxic.

OS Robinia pseudoacacia.

PN WO9911278-A1.  
XX 11-MAR-1999.  
XX 28-AUG-1998; 98WO-GB02612.  
XX 29-AUG-1997; 97GB-0018413.  
XX (ALIZ-) ALIZYME THERAPEUTICS LTD.  
XX Bardocz ZM, Fish NW, Palmer RMJ, Pusztai AJ;  
XX WPI; 1999-228935/19.  
XX Use of Robinia pseudoacacia lectin in medicine  
XX Disclosure; Fig 2; 37pp; English.  
XX The invention relates to Robinia pseudoacacia lectins which can be used  
XX in medicine. The lectin is used for the control of mucosal cell  
XX proliferation, for the reduction and/or treatment of damage caused by  
XX cell damaging agents, especially in the treatment of cancer, and/or for  
XX the reduction and/or treatment of metabolic disorders. It is especially  
XX useful for the treatment of mucositis in mammalian cells and/or tissues,  
XX particularly human cells and/or tissues (especially mucous cells  
XX including mucous membrane). Use of the lectin is effective and does not  
XX damage the gut as is the case with cytotoxic drugs and radiation. The  
XX present sequence represents a R. pseudoacacia lectin RPBAl. This lectin  
XX is composed of two subunits- polypeptide a and polypeptide b  
XX (AA06811-12).  
SQ Sequence 285 AA;  
Query Match 30.6%; Score 489.5; DB 20; Length 285;  
Best Local Similarity 43.6%; Pred. No. 2.2e-37;  
Matches 113; Conservative 37; Mismatches 86; Indels 23; Gaps 8;  
QY 3 SLSEFNKFDLDKQLIFQGDATSTN-NVLQLTKLDGSGNPVGASVGRVLSAPFHLWEN 61  
Db 34 SLSEFPKFPAPNOPYLIFQRDALVTS'GVQLTNN-VNGVPPRRSIRGALYAAPFQIWDN 92  
QY 62 SMA-VSSFETNLTIQISTPHPIYAADGFAFFLAPHDTVIPPNMSGKFLGLYSNVFRNSPT 120  
Db 93 TTGNVASFVTSFSEFIQAPNPATTAADGLAFAPVDT--QPGDLGGMLGIFKD----- 143  
QY 121 SENQSGDVNTDSRVVAVEFDTFPNANIDPNYRHIGIDVNSIKSKETARWQNGKTATA 180  
Db 144 -----GSYNKSNQIVAVEFDTFSNIHFDPKGRHMGINVNSIVSKTVPNWNTNCEVANV 197  
QY 181 RISYNSAKSKSTVTTPYPGMEVVALSHD-VDLHAELPEWVRVGLSASTGEEK---QKNTI 236  
Db 198 FISYEASTKSLNASLVPSLETSTFIHAIVDVKVLPENWVRVGFSAITGIDTYVQTNV 257  
QY 237 ISWSTSSLK-NNEVKEPK 254  
Db 258 LSWSFESNLPGGNSVASVK 276  
RESULT 11  
AA06812  
ID AAY06812 standard; protein; 285 AA.  
XX AAY06812;  
AC AAY06812;  
XX 23-JUN-1999 (first entry)  
DT 23-JUN-1999 (first entry)  
XX R. pseudoacacia lectin RPBAl polypeptide subunit b.  
DE Robinia pseudoacacia; lectin; medicine; mucosal cell proliferation;  
XX Robinia pseudoacacia; lectin; medicine; mucosal cell proliferation;  
KW cancer; metabolic disorder; mucositis; cytotoxic.  
XX cancer; metabolic disorder; mucositis; cytotoxic.  
OS Robinia pseudoacacia.  
XX Robinia pseudoacacia.  
PN WO9911278-A1.

PN WO9911278-A1.  
XX 11-MAR-1999.  
XX 28-AUG-1998; 98WO-GB02612.  
XX 29-AUG-1997; 97GB-0018413.  
XX (ALIZ-) ALIZYME THERAPEUTICS LTD.  
XX Bardocz ZM, Fish NW, Palmer RMJ, Pusztai AJ;  
XX WPI; 1999-228935/19.  
XX Use of Robinia pseudoacacia lectin in medicine  
XX Disclosure; Fig 1; 37pp; English.  
XX The invention relates to Robinia pseudoacacia lectins which can be used  
XX in medicine. The lectin is used for the control of mucosal cell  
XX proliferation, for the reduction and/or treatment of damage caused by  
XX cell damaging agents, especially in the treatment of cancer, and/or for  
XX the reduction and/or treatment of metabolic disorders. It is especially  
XX useful for the treatment of mucositis in mammalian cells and/or tissues,  
XX particularly human cells and/or tissues (especially mucous cells  
XX including mucous membrane). Use of the lectin is effective and does not  
XX damage the gut as is the case with cytotoxic drugs and radiation. The  
XX present sequence represents a polypeptide subunit of R. pseudoacacia  
XX lectin RPBAl.  
SQ Sequence 285 AA;  
Query Match 30.4%; Score 485.5; DB 20; Length 285;  
Best Local Similarity 44.8%; Pred. No. 5.1e-37;  
Matches 116; Conservative 39; Mismatches 81; Indels 23; Gaps 9;  
QY 3 SLSEFNKFDLDKQLIFQGDATSTN-NVLQLTKLDGSGNPVGASVGRVLSAPFHLWEN 61  
Db 34 SLSEFPKFKHSQPDILFQSDALVTSKGVQLTIVNDG--RVYDSIGRVLAAAFQIWD 91  
QY 62 SMA-VSSFETNLTIQISTPHPIYAADGFAFFLAPHDTVIPPNMSGKFLGLYSNVFRNSPT 120  
Db 92 TTGNVASFVTSFSEFIKAPNEGKTADGLVFLAPVGST-QPLKGGGLGLF----- 141  
QY 121 SENQSGDVNTDSRVVAVEFDTFPNANIDPNYRHIGIDVNSIKSKETARWQNGKTATA 180  
Db 142 -KDESY--NKSQIVAVEFDTFERNVAVDENGIMGIDVNSIQSVTRVWDWANGEVANV 197  
QY 181 RISYNSAKSKSTVTTPYPGME-VVALSHDVLHAELPEWVRVGLSASTG---EEKQNTI 236  
Db 198 FISYEASTKSLTASLVPSLEKSFILSAIVDLKKVLPENWVRVGFATTGLSEYVQTNV 257  
QY 237 ISWSTSSLK-NNEVKEPK 254  
Db 258 LSWSFESNLPGGNSVASVK 276  
RESULT 12  
AA06815  
ID AAY06815 standard; protein; 285 AA.  
XX AAY06815;  
AC AAY06815;  
XX 23-JUN-1999 (first entry)  
DT 23-JUN-1999 (first entry)  
XX Amino acid sequence of lectin RPBAlI from R. pseudoacacia seed.  
DE Robinia pseudoacacia; lectin; medicine; mucosal cell proliferation;  
XX Robinia pseudoacacia; lectin; medicine; mucosal cell proliferation;  
KW cancer; metabolic disorder; mucositis; cytotoxic.  
XX cancer; metabolic disorder; mucositis; cytotoxic.  
OS Robinia pseudoacacia.  
XX Robinia pseudoacacia.  
PN WO9911278-A1.

XX PD 11-MAR-1999.  
 XX PF 28-AUG-1998; 98WO-GB02612.  
 XX PR 29-AUG-1997; 97GB-0018413.  
 XX PA (ALIZ-) ALIZYME THERAPEUTICS LTD.  
 XX PI Bardocz ZM, Fish NW, Palmer RMJ, Pusztai AJ;  
 XX WPI; 1999-228935/19.

XX Use of Robina pseudoacacia lectin in medicine

XX Disclosure; Fig 2; 37pp; English.

XX The invention relates to Robinia pseudoacacia lectins which can be used  
 CC in medicine. The lectin is used for the control of mucosal cell  
 CC proliferation, for the reduction and/or treatment of damage caused by  
 CC cell damaging agents, especially in the treatment of cancer, and/or for  
 CC the reduction and/or treatment of metabolic disorders. It is especially  
 CC useful for the treatment of mucositis in mammalian cells and/or tissues,  
 CC particularly human cells and/or tissues (especially mucous cells  
 CC including mucous membrane). Use of the lectin is effective and does not  
 CC damage the gut as is the case with cytotoxic drugs and radiation. The  
 CC present sequence represents a R. pseudoacacia lectin RPBII. This lectin  
 CC is composed of the subunit- polypeptide c (AAY06813).

XX Sequence 285 AA;

Query Match 30.2%; Score 483.5; DB 20; Length 285;

Best Local Similarity 44.0%; Pred. No. 7.9e-37;

Matches 114; Conservative 37; Mismatches 85; Indels 23; Gaps 8;

QY 3 SLSEFTFDLDQKDLIFQGDATSTN-NVLQTLKLDGSGNPVGASVGRVLFSAPEHLWEN 61  
 DB 34 SLSEFTFPAPNPQYLIFQDALVTSTGVQLTNN-VNGVPSKSLGRALYAPFQIWD 92  
 QY 62 SMA-VSSFTNLTIQISTPHYPYAADGFAFFLAPHDTVIPPNWSKGLGLYSNVFRNSPT 120  
 DB 93 TTGNVASEVTSFSLIQAPNPATTADGLAFLAPVDVT--QPLDLGGMGLGIFKN----- 143  
 QY 121 SENQSGDVNTDSRVAVVEFDTPFNANIDPNYRHIGIDVNSIKSKETARWONGKATA 180  
 DB 144 -----GYFNKSNQIVAVEFDTFNRHWDPTGRHMGINYSIVSVKTVPNWANGEVANV 197  
 QY 181 RITSYNSASKKSTVTTFYPGMEVVALSHD-VDLHAELPEWVRVGLSASTGEK---OKNTI 236  
 DB 198 FTSYASTKSLTASLYPSLETSTFIHAIVDVKDVLPEWVRFGFSATTGIDTGYVQNDV 257

QY 237 ISWFTSSLK-NNEYKEPK 254

DB 258 LSWSPESNLPGGNSVASVK 276

RESULT 13

AA58738

ID AAY58738 standard; Protein; 275 AA.

XX AC AAY58738;

XX DT 25-APR-2000 (first entry)

XX DE Pea lectin.

XX KW Lectin; pea; insecticide; transgenic plant; Brassica;

XX KW insect resistance.

XX OS Pisum sativum.

XX PN WO20000123-A1.

XX XX

PD 13-JAN-2000.  
 XX PF 02-JUL-1999; 99WO-SE01209.  
 XX PR 03-JUL-1998; 98SE-0002425.  
 XX PA (PLAN-) PLANT SCI SVERIGE AB.  
 XX PI Ahman I, Melander M, Vamling K;  
 XX WPI; 2000-160693/14.  
 XX N-PSDB; AA258019.

XX Novel lectins used to produce transgenic Brassica plants which are  
 PT resistant to insects

XX Example 2b; Fig 9; 51pp; English.

XX The present sequence is that of pea cv. Lincoln lectin. The  
 CC invention relates to: a transformed Brassica plant that is  
 CC resistant to certain insect pests; an expression cassette  
 CC containing DNA that codes for at least 1 lectin selected from Con A  
 CC (see AAY58736), modified Con A (see AAY58737) and pea lectin;  
 CC transgenic plant cells containing at least 1 copy of the DNA; a new  
 CC lectin derived from the Jack bean Con A gene; a method of imparting  
 CC resistance to insects selected from blossom beetles (pollen  
 CC beetles) of the genus Meligethes, flea beetles of the genus  
 CC Phyllotreta, and root flies of the genus Delia; and a method for  
 CC protecting a plant against infestation by insects of these genera.

XX Sequence 275 AA;

Query Match 29.5%; Score 471; DB 21; Length 275;

Best Local Similarity 40.5%; Pred. No. 1.1e-35;

Matches 107; Conservative 42; Mismatches 85; Indels 30; Gaps 6;

QY 2 QLSFNTKFDLDQKDLIFQGDATSTNNVLQTLKLDGSGNPVGASVGRVLFSAPEHLWEN 61  
 DB 32 ETTSFLTKRTPQDOONLIFQGDGYTTREKLTLC-----AVKNTVGRALYSSPFIHWDR 85  
 QY 62 SMA-VSSFTNLTIQISTPHYPYAADGFAFFLAPHDTVIPPNWSKGLGLYSNVFRNSPT 120  
 DB 86 ETGNVAVNFVSFTVINAPNSYNVADGFTFFIAPVDVT--KPQTGGGVLGVFNS----- 136  
 QY 121 SENQSGDVNTDSRVAVVEFDTPFNANIDPNY--RHIGIDVNSIKSKETARWONGKATA 178  
 DB 137 -----AEYDKTTQTVAVEFDTFYNAAMDPSNRDRHIGIDVNSIKSKVNTKSKLQNGEEA 190  
 QY 179 TARTSYNSASKKSTVTTFYP-----GMEVVALSHDVLHAELPEWVRVGLSASTGEKQ 232  
 DB 191 NVVTAFAATNVLTVSLTYPNSLEEENVTSTLSDVSLKDVVPEWVRFGFSATTGAEYA 250  
 QY 233 KNTIISWTSSTSLKNNVEKPEK 256  
 DB 251 AHEVLWSFHSLSGTSSTSSKQAAD 274

RESULT 14

AAR74766

ID AAR74766 standard; protein; 228 AA.

XX AC AAR74766;

XX DT 13-OCT-1995 (first entry)

XX DE Pea lectin, 2ltn.

XX KW Pentraxin; legume; lectin; structural homology; serum amyloid P; SAP;

XX KW amyloid; fibril; human; Limulus; C-reactive protein; CRP; pea;

XX KW female hamster protein; 3-D structure; jellyroll topology;

XX KW picornavirus; coat protein; pentameric structure; concanavaline A.

XX OS Pisum sativum.

XX WO9505394-A.  
 XX 23-FEB-1995.  
 XX 17-AUG-1994; 94WO-GB01802.  
 XX 17-AUG-1993; 93GB-0017120.  
 XX (BIRK-) BIRKBECK COLLEGE.  
 XX (ROYA-) ROYAL POSTGRAD MED SCHOOL.  
 XX Blundell TL, Pepys MB;  
 XX WPI; 1995-098720/13.  
 XX New cpds. Inhibiting binding of serum amyloid P to amyloid  
 XX fibrils - produced by computer assisted molecular design, useful  
 XX for preventing, treating or diagnosing amyloidosis, e.g.  
 XX Alzheimer's disease  
 XX Claim 7; Fig 4; 72pp; English.  
 XX The sequences given in AAR74763-70 represent various pentraxins and  
 XX legume lectins. All these proteins show structural homology and may  
 XX be used in the molecular design of a molecule for the inhibition of  
 XX serum amyloid P (SAP) binding to amyloid fibrils. The similarities  
 XX in the amino acid sequences of SAP, human and Limulus C-reactive  
 XX protein (CRP) and female hamster protein suggests that they may have  
 XX similar 3-D structures. The jellyroll topology of the pentraxins is  
 XX reminiscent of the picornavirus coat proteins which also have  
 XX pentameric structures. However, pentaxins resemble more closely  
 XX legume lectins such as concanavaline A and pea lectin. Alignment  
 XX of sequences on the basis of topologically equivalent features of the  
 XX three dimensional structures shows that helices occupy different  
 XX positions in the pentraxins and legume lectins and that the amino acid  
 XX sequences of the two families have identities of only approx. 11%. The  
 XX two main helices in SAP occur before and after strand L, whereas the  
 XX helices in the legume lectins occur at the C-terminus of strand J.  
 XX There is a long insertion between the end of the helix after strand D  
 XX and the beginning of strand E in the lectins relative to pentraxins.  
 XX Strands G, H and I together with the type IV beta-hairpin between H and  
 XX I are identical in both SAP and pea lectin. The so-called pentraxin  
 XX octapeptide signature sequence, HXCXS/TWXS, is in this region so it is  
 XX not conserved in the legume lectins.  
 XX Sequence 228 AA;  
 Query Match 26.0%; Score 416; DB 16; Length 228;  
 Best Local Similarity 41.3%; Pred. No. 1.2e-30;  
 Matches 92; Conservative 38; Mismatches 69; Indels 24; Gaps 5;  
 QY 26 STNNVQLTKLDSGGNPVGASVGRVLFSAFPHLWNSMA-VSSPETNLTIQISPHPYA 84  
 Db 2 TTEKELTLK-----AKVTVGRALYSFIIHMDRTGNVANEVTSFTFVLNAPNSVNV 55  
 QY 85 ADGFAFLAPHDVIPPNSWGKFLGLYSNVFNPSPTSENQSGDVNTDSRVVAVEEDTFP 144  
 Db 56 ADGFTFFIAPVDV--KPQTGGVGLVFN-----AEYDKTQTVAVEEDTFY 100  
 QY 145 NANIDPNY--RHIGIDVNSIKSETARWQNGKTATARYNSASKSVITTFYPGMEV 202  
 Db 101 NAAMDVSNRDRHIGIDVNSIKSVNTKSWKLONGEEANVIAFNAATNVLTVSLTYPNVTS 160  
 QY 203 VALSHVDVLHAEFLPEWVRVGLSASTGEKQKNTIISNFTSSL 245  
 Db 161 YTLSDVVSLKDVPEWVRIGFSATTGAEYAHAHEVLSNFSHSEL 203  
 RESULT 15  
 AAY27265  
 ID AAY27265 standard; protein; 262 AA.  
 XX

AC AAY27265;  
 XX 22-OCT-1999 (first entry)  
 XX B. purpurea native lectin sequence.  
 XX Bauhinia purpurea lectin; lectin; BPL; larvicide; insect control;  
 XX European corn borer; Ostrinia nubilalis; crop protection;  
 XX insect larvae infestation.  
 XX Bauhinia purpurea.  
 XX US5945589-A.  
 XX 31-AUG-1999.  
 XX 24-MAR-1993; 93US-0038761.  
 XX 24-MAR-1993; 93US-0038761.  
 XX 24-JUL-1992; 92US-0921179.  
 XX (PION-) PIONEER HI-BRED INT INC.  
 XX Balasubramaniam NK, Rao AG;  
 XX WPI; 1999-508204/42.  
 XX Derivatives of Bauhinia purpurea lectins useful as larvicides to  
 XX control insects (e.g. European corn borer) in field crops such as  
 XX corn, wheat and rapeseed  
 XX Example 5; Columns 11-14; 8pp; English.  
 XX The invention describes compounds derived from Bauhinia purpurea lectin  
 XX ((BPL), a 262 amino acid peptide comprising 7 lysine residues - the  
 XX present sequence) by the replacement of one or more lysine residues with  
 XX other amino acid residues which either preserve the positive charge or  
 XX provide a neutral residue at the point of substitution. The compounds are  
 XX useful as larvicides against insects such as the European corn borer  
 XX (Ostrinia nubilalis). The compounds may be used to protect commercial and  
 XX field crops from insect larvae infestation, especially by the European  
 XX corn borer. The BPL-derivatives are effective against insects which  
 XX may be resistant to the naturally occurring form of BPL. Crops which may  
 XX be protected using the BPL derivatives are corn/maize (Zea mays), rye,  
 XX barley, wheat (Triticum aestivum), sorghum (Sorghum bicolor), oats,  
 XX millet, rice, triticale, sunflower, alfalfa, rapeseed and soybean), fiber  
 XX crops (such as cotton), fruit crops (such as melons) and vegetable crops  
 XX (such as onion, pepper, tomato, cucumber, squash, carrot, crucifer (e.g.  
 XX cabbage, broccoli and cauliflower), eggplant, spinach, potato and  
 XX lettuce). The compounds are useful for protecting species from the genera  
 XX Fragaria, Lotus, Medicago, Onobrychis, Trifolium, Trigonella, Vigna,  
 XX Citrus, Linum, Geranium, Manicot, Daucus, Arabidopsis, Brassica,  
 XX Raphanus, Sinapis, Atropa, Capsicum, Datura, Hyoscyamus, Lycopersicon,  
 XX Nicotiana, Solanum, Petunia, Digitalis, Majorana, Cichorium, Helianthus,  
 XX Lactuca, Bromus, Asparagus, Antirrhinum, Hemerocallis, Nemesis,  
 XX Pelargonium, Panicum, Pennisetum, Ranunculus, Senecio, Salpiglossis,  
 XX Cucumis, Browallia, Glycine, Lolium, Triticum and datura. The present  
 XX sequence represents the B. purpurea native lectin sequence.  
 XX Sequence 262 AA;  
 Query Match 24.0%; Score 383; DB 20; Length 262;  
 Best Local Similarity 40.3%; Pred. No. 1.8e-27;  
 Matches 96; Conservative 39; Mismatches 85; Indels 18; Gaps 7;  
 QY 17 DLIFOGDATSTNNVQLTKLDSGGNPVGASVGRVLFSAFPHLWNSMAVSFETNLTIQI 76  
 Db 23 ELIFLGNATYTPGALRTIGEDGIPKSNAGQASVSRPVFLMDSTGHVASFTSFSFV 82  
 QY 77 -STPHPYAADGFAFLAPHDVIPPNSWGKFLGLYSNVFNPSPTSENQSGDVNTDSRV 135  
 Db 83 RSIDVPHITADGFAFLAPVDSSV--KDYGGCLGLFRYKTATDP-SKNQ-----V 129

1193 VTTYPGMEVVALSHDVLHAELPEWVRVGLSASTGEKQKNTIISWSFTSSLKNEV 250  
| | ||||| ||||| : ||||| :  
1190 VLLTYDNGRHYOLSHVVDPKILPERVRIGFSGGTG-FNETOYILSWSFSTLNSTKI 246

```

completed: February 26, 2003, 16:43:19
time : 40.0712 secs

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:43:23 ; Search time 28.7491 Seconds  
(without alignments)  
2171.628 Million cell updates/sec

Title: US-09-476-485A-6

Perfect score: 1599

Sequence: 1 AQSLSFNFTKFDLDQKDLIF.....LNNHKYVRCSTCMLFMKKK 303

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL\_21.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_ordanelle.\*
  - 9: sp\_phage.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*
  - 15: sp\_rvirus.\*
  - 16: sp\_bacteriap.\*
  - 17: sp\_archaeap.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	1464	91.6	279	10 Q9M7M4	Q9m7m4 phaseolus v
2	804	50.3	272	10 Q9ZTA9	Q9zta9 dolichos la
3	658	41.2	290	10 Q947H0	Q947h0 canavalia e
4	656	41.0	290	10 Q04672	O04672 canavalia b
5	589.5	36.9	266	10 Q93536	P93536 sophora jap
6	582.5	36.4	284	10 Q9FYU9	Q9fyu9 sophora fla
7	579.5	36.2	293	10 P93537	P93537 sophora jap
8	513	32.1	286	10 P93248	P93248 maackia amu
9	510	31.9	282	10 Q93247	Q93247 maackia amu
10	504	31.5	254	10 Q43376	Q43376 arachis hyp
11	498.5	31.2	285	10 Q9ZWP6	Q9zwp6 robinia pse
12	489	30.6	280	10 Q43374	Q43374 arachis hyp
13	481	30.1	254	10 Q43377	Q43377 arachis hyp
14	477	29.8	279	10 Q49899	O49899 medicago sa
15	471	29.5	251	10 Q93X48	Q93x48 lens ervoid
16	467	29.2	251	10 Q8RW33	Q8rw33 lathyrus sa

17	465	29.1	251	10 Q93X42	Q93x42 lens nigric
18	465	29.1	251	10 Q93WH6	Q93wh6 lens culina
19	464	29.0	251	10 Q93X41	Q93x41 lens odemen
20	464	29.0	251	10 Q8W4Y4	Q8w4y4 lens lamott
21	462	28.9	251	10 Q8VXF2	Q8vxf2 lens culina
22	461	28.8	251	10 Q93X50	Q93x50 lens culina
23	461	28.8	251	10 Q93X49	Q93x49 lens culina
24	457.5	28.6	275	10 Q8RW23	Q8rw23 phaseolus c
25	456	28.5	247	10 Q8W1R6	Q8w1r6 griffonia s
26	456	28.5	258	10 Q9FVF8	Q9fvf8 ulex europe
27	456	28.5	275	10 Q8RVY1	Q8rvy1 phaseolus v
28	455.5	28.5	251	10 Q8W1R7	Q8w1r7 griffonia s
29	452.5	28.3	256	10 P93246	P93246 maackia amu
30	451.5	28.2	273	10 Q8RVY4	Q8rvy4 phaseolus c
31	446.5	27.9	258	10 Q41263	Q41263 griffonia s
32	444.5	27.8	276	10 Q40750	Q40750 phaseolus a
33	439.5	27.5	275	10 Q8RVH1	Q8rvh1 phaseolus v
34	438.5	27.4	275	10 Q8RVH3	Q8rvh3 phaseolus v
35	437.5	27.4	275	10 Q8RVY0	Q8rvy0 phaseolus v
36	437.5	27.4	275	10 Q8RVX6	Q8rvx6 phaseolus v
37	435.5	27.2	273	10 Q8RVH2	Q8rvh2 phaseolus v
38	435.5	27.2	278	10 Q9LED8	Q9led8 phaseolus l
39	433.5	27.1	273	10 Q8RVX9	Q8rvx9 phaseolus v
40	430.5	26.9	273	10 Q42411	Q42411 medicago sa
41	430.5	26.9	278	10 Q9LED9	Q9led9 phaseolus l
42	430.5	26.9	285	10 Q9ZWP5	Q9zwp5 robinia pse
43	430	26.9	274	10 Q43628	Q43628 phaseolus v
44	429	26.8	259	10 Q8RVW9	Q8rvw9 vicia faba
45	428.5	26.8	268	10 P93458	P93458 phaseolus l

## ALIGNMENTS

RESULT 1

Q9M7M4 ID Q9M7M4 PRELIMINARY; PRT; 279 AA.  
AC Q9M7M4;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Mannose lectin FRIL (Fragment).  
OS Phaseolus vulgaris (Kidney bean) (French bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.  
OX NCBI\_TaxID=3885;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Moore J.G., Colucci G., Fuchs C.A., Hicklin D.J., Chrispeels M.J.,  
RA Feldman M.;  
RT "A new lectin in red kidney bean called PvFRIL stimulates  
RT proliferation of NIH3T3 cells expressing the Flt3 receptor."  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF121458; AAF28739.1;  
DR HSSP; P02866; 1ONA.  
DR InterPro; IPR000985; Lectin\_legA.  
DR InterPro; IPR001220; Lectin\_legB.  
DR Pfam; PF00138; lectin\_legA; 1.  
DR Pfam; PF00139; lectin\_legB; 1.  
DR PRODOM; PD000671; Lectin\_legA; 1.  
DR PRODOM; PD000711; Lectin\_legB; 1.  
DR PROSITE; PS00308; LECTIN\_LEGUME\_ALPHA; 1.  
DR PROSITE; PS00307; LECTIN\_LEGUME\_BETA; UNKNOWN\_1.  
KW Lectin.  
FT NON-TER.  
SQ SEQUENCE 279 AA; 31102 MW; F8919CF8B3EE4652 CRC64;

Query Match 91.6%; Score 1464; DB 10; Length 279;  
Best Local Similarity 100.0%; Pred. No. 3.6e-106;  
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AQSLSFNFTKFDLDQKDLIFQGDATSTNNVQLTKLDGGNPVGASVGRVLSAPFHLWE 60

Qy	241	FTSSLKNNVEKPKEDMYIANVRSYTWINDVLSYISNK	279
Db	241	FTSSLKNNVEKPKEDMYIANVRSYTWINDVLSYISNK	279

GN	CONA.	
OS	<i>Canavalia ensiformis</i> (Jack bean) (Horse bean).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;	
OC	eurosid1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Canavalia	
ON	NCBI_TaxID=3823;	
OX	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN-CV. U-02;	
RA	Ramis C.; Gomord V.;	
RL	Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF308777; AAL09432.1; -	
DR	InterPro; IPR000985; Lectin_legA.	
DR	InterPro; IPR001220; Lectin_legB.	
DR	Pfam; PF001138; lectin_legA; 1.	
DR	Pfam; PF001139; lectin_legB; 1.	
DR	ProDom; PD000671; lectin_legA; 1.	
DR	ProDom; PD000711; lectin_legB; 1.	
DR	PROSITE; PS00308; LECTIN_LEGUME_ALPHA; UNKNOWN_1.	
DR	PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.	
DR	SEQUENCE 290 AA; 31480 MW; 0F2F7DBRCF547E42 CRC64;	
SO		

	Query Match	41.2%	Score 658;	DB 10;	Length 290;
	Best Local Similarity	52.2%;	Pred. No. 2.le-43;		
	Matches 140;	Conservative 42;	Mismatches 66;	Indels 20;	Gaps
Qy	3	SLSNFETKFDLDQKDLIFQGDSATS-TNNVLVLTKLDSGGNPVGASGVRVLSAFPFLHWN 61	:	:	:
Db	36	ALHFMEHQFSKDQKDLILQGDATTGTGNLELTRVSSNGSPQSGSVGRALFYAPVHIWES 95	:	:	:
Qy	62	SMAYSSPETNLTIQISTP--HPYYAADGFAFFLAPHDTVIPPNWSMGKFGLGY--SNVFRN 11	:	:	:
Db	96	SAVVASEFATEFLIKSDPSHP---ACGAFFINIDISSIPSGSTGRLGLLPDANVIRN 15	:	:	:
Qy	118	SPTSE-NQSFGDVNTDSRVAAVEEDTFPNANI-DPNYRHIGIDIVNSKSKETAREWQNG 17	:	:	:
Db	153	STTIDFNAAY--NADT-IVAVELDTPTNDIGDPSPYHIGIDIKSVRSKKATAKNMQNG 20	:	:	:
Qy	176	KTATARTSYNSASKKSVTPTYTPGMENVVALSHVDLLHAELPEWTVRVLGSASTGEOKNT 23	:	:	:
Db	209	KVGTAHIIYSVDRKLSAVSYPNADSATSYDVDLDNVLPPEWRVRLGSASTGLYKETNT 26	:	:	:
Qy	236	IISWFTSSLKNNVEKPKEDMYIANVY 263	:	:	:
Db	269	ILSWFTSKLKSNEIPD-----IATVV 290	:	:	:
RESULT 4					
O04672	ID	PRELIMINARY;	PRT;	290 AA.	
AC	O04672;				
DT	01-JUL-1997 (TREMBLrel. 04, Created)				
DT	01-JUL-1997 (TREMBLrel. 04, Last sequence update)				
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)				
DE	Lectin (fragment).				
OS	Canavalia brasiliensis (Brazilian jack bean).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae				







P		SEQUENCE FROM N.A.	
P		yoshida K., Tazaki K.:	
T		"Expression patterns of the genes that encode lectin or lectin-related	
A		polypeptides in Robinia pseudoacacia."	
L		Aust. J. Plant Physiol. 26:495-502(1999).	
R		EMBL; AB012633; BAA36414.1; -	
R		HSSP; P19588; 1LUL	
R		InterPro: IPR000985; Lectin_legA.	
R		InterPro: IPR001220; Lectin_legB.	
R		Fram; PF00138; lectin_legA; 1.	
R		Fram; PF00139; lectin_legB; 1.	
R		ProDom; PD000671; Lectin_legA; 1.	
R		ProDom; PD000711; Lectin_legB; 1.	
R		PROSITE; PS00308; LECTIN_LEGUME_ALPHA; UNKNOWN_1.	
R		PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.	
SQ		SEQUENCE 285 AA; 30940 MW; 5B5A42C8B9579922 CRC64;	
		Query Match	31.2%; Score 498.5; DB 10; Length 285;
		Best Local Similarity	44.8%; Pred. No. 5.3e-31;
		Matches 116; Conservative	36; Mismatches 84; Indels 23; Gaps
QY		3 SLSEFNFTKFDLDQKDLPDGDATSTN-NVLQLTKLDSCGNPGVSGRVLSAPFLWEN 61	
Ddb		:           :           :           :           :           :	
		34 SLSEFPFRFAPNQPYLLIQGDALVTGTGLOLTNV-VNGVFSRKSGLRALYAAPFIWDS 92	
QY		:           :           :           :           :           :	
Ddb		62 SMA-VSSETNLTIQISHPHYAADGAFFLAHDTVIPNSMGKFGLGSYNFERNST 120	
QY		:           :           :           :           :           :	
Ddb		93 TTGNVGASVFESFTIIQAPNPATTADGLAFPLAVDT--QPDLIGMLGIPTKN----- 143	
QY		121 SENQSGVDNTSDRVVAVEFDTPPNANIDPNRYHIGIDVNSIKSKETARWEHQNKATA 180	
QY		:       :           :           :           :           :	
Ddb		144 -----GVFNKSNQIVAVEFDTFSRNHMDPTGRHLGINVINSIKSVRTVPWNNTGEVANV 197	
QY		181 RIYSNASAKSKNTTYFGMEVVALLSHD-VDLHAELPEWVRVGLSASTGEEK---QKNTI 236	
Ddb		198 FTSYEASTKSLTASLVPSLETSFVIAHVVDVKDLPEWVRFGSATTTGIDKGYYQTNDV 257	
QY		237 ISWSETSSLK-NNEVKEPK 254	
Ddb		258 LSWSFESNLPGGNSVASVK 276	
		:	
RESULT 12			
Q43374		PRELIMINARY;	PRT; 280 AA.
ID	Q43374		
AC	Q43374;		
DT	01-NOV-1996 (TrEMBLrel. 01, Created)		
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)		
DE	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	Mannose/glucose-binding lectin precursor.		
GN	LEC.		
OS	Arachis hypogaea (Peanut),		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eurosid1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;		
OX	Arachis.		
OX	NCBI_TaxId=3818;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-CV. SELLIE; TISSUE-MODULE;		
KT	Law I.J.;		
KT	"Cloning and expression of cDNA for mannose-binding lectin from		
RL	peanut."		
RL	Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.		
DR	EMBL; U22469; AAA74572.1; -		
DR	HSSP; P02866; 1DQ2		
DR	InterPro: IPR000985; Lectin_legA.		
DR	InterPro: IPR001220; Lectin_legB.		
DR	Fram; PF00138; lectin_legA; 1.		
DR	Fram; PF00139; lectin_legB; 1.		
DR	ProDom; PD000671; Lectin_legA; 1.		
DR	ProDom; PD000711; Lectin_legB; 1.		
DR	PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.		

```

KW Lectin; Signal.
FT SIGNAL 1 26
FT CHAIN 27 280 MANNOSYLGLUCOSE-BINDING LECTIN.
SQ SEQUENCE 280 AA; 31012 MW; C15B39B32F455BD5 CRC64;

Query Match
Best Local Similarity 30.6%; Score 489; DB 10; Length 280;
Matches 111; Conservative 45; Mismatches 75; Indels 18; Gaps 7;

QY 3 SLSFNFTKFDL-DOKDLIFOGDAT-STNNVQLTKLDSGPNVGSVGRVLSAPFHLWE 60
DB 29 SLSFSYNNEFQDDERNLILQGDATSAKGIQLTKVDNDGTPAKSTVGRVLSHTQVRLWE 88
QY 61 NSM-AVSSEFTNLTIQISTPHPYAAGDAFFLPHADTVIPNSWGKFLGLYSNVRNSP 119
DB 89 KSTNRLTNFOAQSFVKSIPID-NGADGIAFFIAAPDSEIPKNSAGGTGLGF-----DPS 142
QY 120 TSENQSGDVNTDSRVVAVEFTF---PNANIDPNYRHIGIDVNSIKSKETARWQNGK 176
DB 143 TAOQPS-----ANQVLAVEFTFYAODSNGWDPNYQHIGIDVNSIKSAATTKWERRNGQ 196
QY 177 TATARISYNSASKKSTVTTFYPGMEVVALSHDVLHAEIPWVRVGLSASTGEKQKNTI 236
DB 197 TLNVLSYDANSKNLQVTASYPDQORYQVSYNVDRLDYLPWGRVGFSAASQQYQSHL 256
QY 237 ISWSFTSSL 245
DB 257 QSWSTSTL 265

RESULT 13
Q43377 PRELIMINARY; PRT; 254 AA.
AC Q43377;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Mannose/glucose-binding lectin precursor (fragment).
GN LEC.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. SELLIE; TISSUE=SEED;
RA Law I.J.;
RT "Cloning and expression of cDNA for mannose/glucose-binding lectin
RT from peanut.";
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U22473; AAA74576.1; -.
DR HSSP; P02867; 2BQP.
DR InterPro; IPR000985; Lectin_legA.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; lectin_legA; 1.
DR ProDom; PD000711; lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_ALPHA; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
KW LECTIN; Signal.
FT SIGNAL 1 2
FT CHAIN 254 AA; 28389 MW; 761882817DE7DFB9 CRC64;
SQ SEQUENCE 254 AA; 28389 MW; 761882817DE7DFB9 CRC64;

Query Match
Best Local Similarity 30.1%; Score 481; DB 10; Length 254;
Matches 110; Conservative 45; Mismatches 76; Indels 18; Gaps 7;

QY 3 SLSFNFTKFDL-DOKDLIFOGDAT-STNNVQLTKLDSGPNVGSVGRVLSAPFHLWE 60
DB 3 SLSFSYNNEFQDDERNLILQGDATSAKGIQLTKVDNDGTPAKSTVGRVLSHTQVRLWE 62

```

```

QY 61 NSM-AVSSEFTNLTIQISTPHPYAAGDAFFLPHADTVIPNSWGKFLGLYSNVRNSP 119
DB 63 KSTNRLTNFOAQSFVKSIPID-NGADGIAFFIAAPDSEIPKNSAGGTGLGF-----DPQ 116
QY 120 TSENQSGDVNTDSRVVAVEFTF---PNANIDPNYRHIGIDVNSIKSKETARWQNGK 176
DB 117 TAOQPS-----ANQVLAVEFTFYAODSNGWDPNYQHIGIDVNSIKSAATTKWERRNGQ 170
QY 177 TATARISYNSASKKSTVTTFYPGMEVVALSHDVLHAEIPWVRVGLSASTGEKQKNTI 236
DB 171 TLNVLSYDANSKNLQVTASYPDQORYQVSYNVDRLDYLPWGRVGFSAASQQYQSHL 230
QY 237 ISWSFTSSL 245
DB 231 QSWSTSTL 239

RESULT 14
O49899 PRELIMINARY; PRT; 279 AA.
AC O49899;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Lactin precursor.
GN LEC2.
OS Medicago sativa (Alfalfa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
OX NCBI_TaxID=3879;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. CHIEF;
RA Brill L.M.; Pieterneel V.R.;
RT "Legume seed lectin genes: sequence of Mslec2 from Alfalfa, Alfalfa
RT and Sweet clover southern blot analysis, and stable transformation of
RT Alfalfa with antisense-lectin constructs.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y16754; CAA76366.1; -.
DR HSSP; P04122; LLOB.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; lectin_legA; 1.
DR ProDom; PD000711; lectin_legB; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
KW SIGNAL.
FT SIGNAL 1 25
FT CHAIN 26 279
FT CHAIN 26 279 LECTIN.
SQ SEQUENCE 279 AA; 30551 MW; 1315F022BABDA360 CRC64;

Query Match
Best Local Similarity 29.8%; Score 477; DB 10; Length 279;
Matches 111; Conservative 36; Mismatches 78; Indels 30; Gaps 6;

QY 1 AQSLSFNFTKFDL-DOKDLIFOGDA-TSTNNVQLTKLDSGPNVGSVGRVLSAPFHLW 59
DB 27 AETTSFSTIKFVPDQKNLIFOGDAETSTGKRLTK-----AVKNSIGRALYSAPIHW 80
QY 60 ENSM-AVSSEFTNLTIQISTPHPYAAGDAFFLPHADTVIPNSWGKFLGLYSNVRN 117
DB 81 DSKTGSVANFETFTTITAPNTYNAVADGLAFIADITOPKNSQGGYLGDFDSTKYQE 140
QY 118 SPTSENQSGDVNTDSRVVAVEFTFNPANIDP-----NYRHIGIDVNSIKSKETARW 170
DB 141 S-----IQTVAVEIDTFYVNDWDPKPGNISSTGRHIGINVSISITVPW 186
QY 171 EWQNGKTATARIYSYNSASKKSTVTTFYPCMEVVALSHDVLHAEIPWVRVGLSASTGEE 230

```

Wed Feb 26 17:04:11 2003

Db 187 SLVNNKANVVGIGATNVLSDVVEYPLVRHYTLSHVVPLKDVVPEWVRVIGFSAATGDE 246  
Qy 231 KOKNTIISWSTSSL 245  
Db 247 YAEHDFSWFSDSKL 261

## RESULT 15

Q93X48 PRELIMINARY; PRT; 251 AA.  
AC Q93X48;  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
DE LECTIN (Fragment).  
GN LECTIN.  
OS Lens ervoides.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Lens.  
OX NCBI\_taxID-41257;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LEAF;  
RA Galasso I.;  
FT "Lectin genes in Lens."  
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ318220; CAC42126.1; -  
DR InterPro; IPR000985; Lectin\_legA.  
DR InterPro; IPR001220; Lectin\_legB.  
DR Pfam; PF00138; lectin\_legA; 1.  
DR Pfam; PF00139; lectin\_legB; 1.  
DR ProDom; PD000671; Lectin\_legA; 1.  
DR ProDom; PD000711; Lectin\_legB; 1.  
DR PROSITE; PS00308; LECTIN\_LEGUME\_ALPHA; UNKNOWN\_1.  
DR PROSITE; PS00307; LECTIN\_LEGUME\_BETA; UNKNOWN\_1.  
FT NON\_TER 1  
FT NON\_TER 251  
SQ SEQUENCE 251 AA; 27963 MW; FE85E9765AF7DFD CRC64;

Query Match 29.5%; Score 471; DB 10; Length 251;  
Best Local Similarity 42.6%; Pred. No. 6.2e-29;  
Matches 106; Conservative 38; Mismatches 75; Indels 30; Gaps 6;  
Qy 2 QSLSFNFTKFDLDKDLIFQGDATSTNNVLQLKLDGSGNPVGASVGRVLFSAFFHLWEN 61  
Db 23 ETTSSITKFSPPQNLIFQGDGYTTKEKLTGK-----AVKNTVGRALYSTPIHIWDR 76  
Qy 62 SMA-VSSFTNLTITQSTHPHYAAGCAFFLAPHDVIPPNSWCKELGLYSNVFRNSPT 120  
Db 77 DTGNVANFVTSFTFVINAPNSYVADGTFPIAPVDT--KPTGGGYLGVFNS----- 127  
Qy 121 SENQSGDVNTDSRVVAVEDTTPNANIDPNY--RHIGIDVNSIKSKETARWQNGKTA 178  
Db 128 -----KDYKTSQTVAVEFTFYNAWDESNKDRHIGIDVNSIKSVKTSWNQNGERA 181  
Qy 179 TARISYNSAKSKSTVTTFYP-----GMEYVALSHVDLHAELEPWEVVRVGLSASTGEKQ 232  
Db 182 NVVIAENATNVLTVLTYPNSLEEENVTSTLNEVVPMKDVLPWVRVIGFSAATGAFA 241  
Qy 233 KNTIISWSF 241  
Db 242 AHEVLSWSF 250

Search completed: February 26, 2003, 16:54:26  
Job time : 29.7491 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:41:33 ; Search time 8.70037 Seconds  
(without alignments)  
1444.458 Million cell updates/sec

Title: US-09-476-485A-6

Perfect score: 1599

Sequence: 1 AQSLSFNFTKFDLDQKDLIF.....LNNHXYVRCSTCMLEPKKK 303

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query No.	Score	Match %	Length	ID	Description
1	658	41.2	290	1	CONA_CANGL	P14894 canavalia g
2	656	41.0	290	1	CONA_CANEN	P02866 canavalia e
3	650	40.7	237	1	LECA_DOLLA	P38662 dolichos la
4	629.5	39.4	240	1	LEC_BOWMI	P42088 bowringia m
5	621	38.8	290	1	LEC2_GIALU	Q39529 cladrastis
6	604.5	37.8	293	1	LEC1_GIALU	Q39528 cladrastis
7	501.5	31.4	290	1	LECR_GIALU	Q39527 cladrastis
8	497	31.1	240	1	LECS_VATMA	P81371 vatireia ma
9	495	31.0	286	1	LEC2_ROBPS	Q42372 robinia pse
10	493.5	30.9	285	1	LEC1_ROBPS	Q41159 robinia pse
11	493	30.8	270	1	LECB_SOPJA	P93538 sophora jap
12	489.5	30.6	285	1	LECS1_ROBPS	Q41162 robinia pse
13	483.5	30.2	285	1	LECS2_ROBPS	Q41161 robinia pse
14	480	30.0	292	1	LECS_SOPJA	P93535 sophora jap
15	478	29.9	280	1	LEC2_WEDIR	Q01807 medicago tr
16	471	29.5	275	1	LEC_PEA	P02867 pisum sativ
17	470	29.4	248	1	LEC2_CYTSC	P29257 cytisus sco
18	459	28.7	273	1	PHAM_PHAVU	P15231 phaseolus v
19	452	28.3	233	1	LEC_VICFA	P02871 vicia faba
20	444.5	27.8	272	1	PHAL_PHAVU	P05087 phaseolus v
21	439	27.5	250	1	LEC1_LABAL	P23358 laburnum al
22	436.5	27.3	249	1	LEC2_ULEEU	P02973 ulex europe
23	435.5	27.2	285	1	LEC_SOYBN	P05046 glycine max
24	433	27.1	275	1	LEC1_DOLBI	P05045 dolichos bi
25	427.5	26.7	275	1	PHAE_PHAVU	P05088 phaseolus v
26	422	26.4	275	1	LEC5_DOLBI	P19588 dolichos bi
27	421.5	26.4	277	1	LEC1_WEDIR	Q01806 medicago tr
28	413	25.8	236	1	LEC_ONOVI	P02874 onobrychis
29	396	24.8	211	1	LEC_LENCU	P02870 lens culina
30	393.5	24.6	243	1	LEC4_GRISI	P24146 griffonia s
31	386	24.1	181	1	LECB_LATOC	P04122 lathyrus oc
32	383	24.0	290	1	LEC_BAUPU	P16030 baubinia pu
33	382.5	23.9	281	1	LEC_ERYCO	P16404 erythrina c

34 380 23.8 244 1 LEC\_LATSP P16349 lathyrus sp  
35 367.5 23.0 242 1 LEC1\_PSOTE O24313 psophocarpu  
36 365.5 22.9 237 1 CONA\_CANLI P81460 canavalia l  
37 364.5 22.8 236 1 LECA\_CRAFL P81517 cratylia fl  
38 363.5 22.7 237 1 CONA\_CANBR P55915 canavalia b  
39 363.5 22.7 237 1 CONA\_CANVI P81461 canavalia v  
40 360.5 22.5 237 1 LECA\_DIOGR P08902 dioclea gra  
41 359.5 22.5 237 1 LECA\_DIOGR P16377 dioclea gra  
42 359 22.5 273 1 LECG\_ARAHY P02872 arachis hyp  
43 351.5 22.0 243 1 LEC1\_ULEEU P22972 ulex europe  
44 350 21.9 265 1 LECN\_PEA P16270 pisum sativ  
45 334 20.9 240 1 LEC\_LOTTE P19664 lotus tetra

## ALIGNMENTS

## RESULT 1

ID CONA\_CANGL STANDARD; PRT: 290 AA.  
AC P14894;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Concanavalin A precursor (Con A).  
OS Canavalia gladiata (Sword bean) (Japanese jack bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Canavalia.  
OX NCBI\_TaxID=3824;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Seedling;  
RX MEDLINE=90127395; PubMed=2404793;  
RA Yamauchi D., Nakamura K., Asahi T., Minamikawa T.;  
RT "Nucleotide sequence of cDNA for concanavalin A from Canavalia  
gladiata seeds".  
RL Plant Cell Physiol. 30:147-150(1989).  
CC -!- FUNCTION: D-MANNOSE SPECIFIC LECTIN.  
CC -!- SUBUNIT: HOMOTETRAMER.  
CC -!- PTM: THE MATURE CHAIN CONSISTS OF RESIDUES 164-281 FOLLOWED BY  
30-148. TO FORM A MATURE CHAIN THE PRECURSOR UNDERGOES FURTHER  
POST-TRANSLATIONAL MODIFICATION AFTER REMOVAL OF THE SIGNAL  
SEQUENCE. CLEAVAGE AFTER ASN AT POSITIONS 148, 163, AND 281 IS  
FOLLOWED BY TRANSPOSITION AND LIGATION (BY FORMATION OF A NEW  
PEPTIDE BOND) OF RESIDUES 164-281 AND 30-148.  
CC -!- MISCELLANEOUS: BINDS ONE MANGANESE (OR OTHER TRANSITION METAL) ION  
AND ONE CALCIUM ION. THE METAL IONS ARE ESSENTIAL FOR THE  
SACCHARIDE-BINDING AND CELL-AGGLUTINATING ACTIVITIES.  
CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.  
-----

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
-----

EMBL; X16041; CAA34163.1; ..  
PIR; A34139; A34139.  
PIR; JQ2130; JQ2130.  
HSP; P02866; IDQ2.  
InterPro; IPR000985; Lectin\_legA.  
InterPro; IPR001220; Lectin\_legB.  
Pfam; PF00138; lectin\_legA; 1.





```

FT METAL 127 127 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 132 132 MANGANESE (BY SIMILARITY).
SQ SEQUENCE 237 AA; 25718 MW; BEE7E84DC2895327 CRC64;

Query Match
Best Local Similarity 40.7%; Score 650; DB 1; Length 237;
Matches 143; Conservative 30; Mismatches 54; Indels 26; Gaps 7;

QY 1 AQSLSFNFTKFDLQKDLIFQGDATSTNNVLQTLKDSGGNPGVGSVGRVLFSAFFHLW 60
  |||||:||||:||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 AQSLSFSFTKFDPNQDLIFQGTATS-----KLDSAGNPVSSAGRVLYSAPLRW 52
  |||||:||||:||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 NSMAVSSFEFTNLTIQISTPHPPYAAADGFAFFLAPHDTVIPPNWKGKFLGLYSNVFRSPT 120
  |||||:||||:||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 53 DSAVLTSFDP--TIYIFTNTSRIADGLA-FIAPPDVSI--SYHGGLGLFPNAESG-- 105
  |||||:||||:||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 SENQSGDGVNTDSRVVAVVEFDTPNPANI-DPNYRHIGIDVNSIKSKETARWQNGKTA 178
  |||||:||||:||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 106 -----IAESNVAVVEFDTPNPYDGNPNYIHIGIDVNSIRSKVTASWDWQNGKIA 156
  |||||:||||:||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 179 TARISYNSASKKSTVTTFPGMEVVALSHDVLHAELPEWVRVGLSASTGEEKKNTIIS 238
  |||||:||||:||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 157 TAHISYNSVSKSLSTVYYPGKPKATSYDIELHTVLPWVRVGLSASTGQNIERTVHS 216
  |||||:||||:||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 239 WSTSSLKNEVK 251
  |||||:||||:||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 217 WSTSSLTWNTVAK 229
  |||||:||||:||||| ||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
LEC_BOWMI STANDARD; PRT; 240 AA.
AC P42088;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Lectin (Agglutinin) (BMA).
OS Bowringia milderadii.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Bowringia.
OX NCBI_TaxID=28956;
RN [1]
RP SEQUENCE.
RC TISSUE=Seed;
RX MEDLINE=93385179; PubMed=8373823;
RA Chawla D., Animeshaun T., Hughes R.C., Harris A., Aitken A.;
RT "Bowringia milderadii agglutinin: polypeptide composition, primary
RT structure and homologues with other legume lectins.";
RL Biochim. Biophys. Acta 1202:38-46(1993).
CC -!- FUNCTION: BINDS PREFERENTIALLY TO OLIGOSACCHARIDES BEARING THE
CC SEQUENCE MAN-ALPHA-1->2 MAN-ALPHA-1->6 MAN-ALPHA-1->6MAN FOUND
CC IN EARLY STEPS OF GLYCOPROTEIN PROCESSING IN THE ENDOPLASMIC
CC RETICULUM. IT BINDS WEAKLY TO HIGHLY PROCESSED OLIGOSACCHARIDE
CC STRUCTURES.
CC -!- COFACTOR: CALCIUM AND MANGANESE ARE ESSENTIAL FOR THE SACCHARIDE-
CC BINDING AND CELL-AGGLUTININATING ACTIVITIES.
CC -!- SUBUNIT: HETEROTETRAMER OF TWO ALPHA AND TWO BETA CHAINS;
CC DISULFIDE BOND LINKED.
CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
CC HSSP; P02866; I0Q2.
DR InterPro: IPR000985; Lectin_legA.
DR InterPro: IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; FALSE_NEG.
KW Lectin; Calcium; Manganese.
FT CHAIN 1 116 BETA CHAIN.
FT CHAIN 117 240 ALPHA CHAIN.
FT METAL 127 127 MANGANESE (BY SIMILARITY).

```

```

FT METAL 129 129 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 131 131 CALCIUM (BY SIMILARITY).
FT METAL 133 133 CALCIUM (BY SIMILARITY).
FT METAL 138 138 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 143 143 MANGANESE (BY SIMILARITY).
FT DISULFID 5 5 INTERCHAIN.
FT VARIANT 143 143 H -> S.
FT VARIANT 158 158 K -> R.
FT VARIANT 177 177 S -> R.
FT VARIANT 187 187 P -> V.
FT VARIANT 203 203 V -> A.
FT UNSURE 86 90
FT UNSURE 116 116
SQ SEQUENCE 240 AA; 25424 MW; 5A9F7FAF3A09B060 CRC64;

Query Match
Best Local Similarity 39.4%; Score 629.5; DB 1; Length 240;
Matches 133; Conservative 33; Mismatches 62; Indels 23; Gaps 5;

QY 1 AQSLSFNFTKFDLQKDLIFQGDAT-STNNVLQTLKDSGGNPGVGSVGRVLFSAFFHLW 59
  |||||:||||:||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 ANSVCFFTDFESGQDLIFQGDASVGNKALQTLKVDKSGNPGQGSVGRALYTAPIRLW 60
  |||||:||||:||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 60 ENSMAVSSFEFTNLTIQI-----STPHPPYAAADGFAFFLAPHDTVIPPNWKGKFLGLYSNVF 115
  |||||:||||:||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 QSSSLVASFEFTTFSISQGSSTP-----AAALTFIASPDTKIPSGGRLGLFG---- 112
  |||||:||||:||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 116 RNSPTSENQSGDVNTDSRVVAVVEFDTPNPANI-DPNYRHIGIDVNSIKSKETARWQNG 174
  |||||:||||:||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 113 -----SSNAGSDGVVAVVEFDTPNTDIGNPNYRHIGIDVNSIRSKAASWDWQNG 163
  |||||:||||:||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 175 GKTATARISYNSASKKSTVTTFPGMEVVALSHDVLHAELPEWVRVGLSASTGEEKQKN 234
  |||||:||||:||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 164 GKTATARISYNSASKKSTVTTFPGMEVVALSHDVLHAELPEWVRVGLSASTGEEKQKN 234
  |||||:||||:||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
QY 235 TISWSFTSSL 245
  |||||:||||:||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 224 NILAWSFRSSL 234
  |||||:||||:||||| ||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
LEC2_CLALU STANDARD; PRT; 290 AA.
AC Q39529;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Agglutinin II precursor (CLAII) (LecCLAII).
OS Cladrastis lutea (Yellow wood).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Cladrastis.
OX NCBI_TaxID=38412;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 36-54.
RC TISSUE=Bark;
RX MEDLINE=96123235; PubMed=8534854;
RA van Damme E.J.M., Barre A., Bemer V., Rouge P., van Leuven F.,
RA Peumans W.J.;
RT "A lectin and a lectin-related protein are the two most prominent
RT proteins in the bark of yellow wood (Cladrastis lutea).";
RL Plant Mol. Biol. 29:579-598(1995).
CC -!- FUNCTION: MANNOSYLGLUCOSE BINDING BARK LECTIN.
CC -!- FUNCTION: BARK LECTINS ARE STORAGE PROTEIN THAT PROBABLY MAINTAINS
CC STOCKS OF NITROGEN DURING DORMANT PERIOD. SELF-AGGREGATABLE
CC MOLECULES THAT CAN BIND THEIR OWN CARBOHYDRATE SIDE CHAINS. THEY
CC COULD ALSO PLAY A ROLE IN THE PLANT'S DEFENCE AGAINST PHYTOPHAGOUS
CC INVERTEBRATES OR HERBIVOROUS HIGHER ANIMALS.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- MISCELLANEOUS: BINDS ONE MANGANESE (OR OTHER TRANSITION METAL) ION
CC AND ONE CALCIUM ION. THE METAL IONS ARE ESSENTIAL FOR THE
CC SACCHARIDE-BINDING AND CELL-AGGLUTININATING ACTIVITIES (BY
CC SIMILARITY).

```



CC	-1- S
DR	HSSP;
DR	Glyco;
DR	Inter
DR	Inter
DR	pfam;
DR	pfam;
DR	ProDo
DR	ProDo
DR	PROSI
DR	PROSI
KW	Lectin
FT	CHAIN
FT	CHAIN
FT	METAL
FT	METAL
FT	METAL
FT	METAL







kw  
Lectin; Calcium; Manganese; Glycoprotein; Signal; Multigene family.

```

DR EMBL; U24249; AAC49271.1; -.
DR HSSP; P19588; 1LUL.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
KW Lectin; Calcium; Manganese; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 31
FT CHAIN 32 285 SEED AGGLUTININ II.
FT METAL 156 156 MANGANESE (BY SIMILARITY).
FT METAL 158 158 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 162 162 CALCIUM (BY SIMILARITY).
FT METAL 166 166 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 171 171 MANGANESE (BY SIMILARITY).
FT CARBOHYD 147 147 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 285 AA; 31021 MW; 2C0B3249620294DE CRC64;

Query Match 30.2%; Score 483.5; DB 1; Length 285;
Best Local Similarity 44.0%; Pred. No. 9.2e-31;
Matches 114; Conservative 37; Mismatches 85; Indels 23; Gaps 8;

QY 3 SLSEFNFTKFDLQKDLIFOGDATSTN-NVLQTLKDSGGNPGVSGRVLFSAPFHLWEN 61
| | | | | : | : | | | : | : | | | : | : | | | : | : | | | : | : |
Db 34 SLSEFPKPKAPNOPIYLFORDALVTSTGVQLTNV-VNGVPSRKSLGRALYAAPFQIWD 92
| | | | | : | : | | | : | : | | | : | : | | | : | : | | | : | : |
QY 62 SMA-VSSFTNLTIQISTPHYYAAGFAFFLPHDTVIPPNWSGKFLGLYSNVFRNSPT 120
| | | | | : | : | | | : | : | | | | | | | | | | | | | | | | | |
Db 93 TTGNVASFDVTSFIIQAPNATPADGLAFPLAPVDT--QPLDGLGMLGIFKN----- 143
| | | | | : | : | | | | | | | | | | | | | | | | | | | | | | |
QY 121 SENQSGDVNTDSRVAVVEFDTPPNANIDPNYRHIGIDVNSIKSKETARWONGKTATA 180
| | | | | : | : | | | | | | | | | | | | | | | | | | | | | | |
Db 144 -----GYENKSNQIVAVEEDFENRHWDPTRGHMGINVSIVKTVPNWANGEVANV 197
| | | | | : | : | | | | | | | | | | | | | | | | | | | | | | |
QY 181 RTSYNSAKSKSTVTTFYPGMEVVALSHD-VDLHAELPEWVRVGLSASTGEEK---QKNTI 236
| | | | | : | : | | | : | : | | | | | | | | | | | | | | | | |
Db 198 FTSYEASTKSLTASLYPSLETFSIIHAIVDKVDLPVWVRFGFSATGIDTGYVQTNDV 257
| | | | | : | : | | | : | : | | | | | | | | | | | | | | | | |
QY 237 ISWFTSSLK-NNEVKPK 254
| | | | | : | : | | | : | : | | | | | | | | | | | | | | | | |
Db 258 LSWSFESNLPGGNSVASVK 276
| | | | | : | : | | | : | : | | | | | | | | | | | | | | | | |

RESULT 14
LECS_SOPJA STANDARD; PRT; 292 AA.
AC P93535;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Seed lectin precursor (LECS7ASG).
OS Sophora japonica (Japanese pagoda tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Sophora.
OX NCBI_TaxID=3897;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=97201486; PubMed=9049272;
RA Van Damme E.J., Barre A., Rouge P., Peumans W.J.;
RT "Molecular cloning of the bark and seed lectins from the Japanese
RT pagoda tree (Sophora japonica).";
RL Plant Mol. Biol. 33:523-536(1997).
CC !- FUNCTION: MANNOSYLGLUCOSE-SPECIFIC LECTIN.
CC !- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U63011; AAB51441.1; -.
DR HSSP; P19588; 1LUL.
DR GlycoSuiteDB; P93535; -.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
KW Lectin; Calcium; Manganese; Glycoprotein; Signal.
FT SIGNAL 1 37 POTENTIAL.
FT CHAIN 38 292 SEED LECTIN.
FT METAL 163 163 MANGANESE (BY SIMILARITY).
FT METAL 165 165 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 167 167 CALCIUM (BY SIMILARITY).
FT METAL 169 169 CALCIUM (BY SIMILARITY).
FT METAL 172 172 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 177 177 MANGANESE (BY SIMILARITY).
FT CARBOHYD 82 82 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 292 AA; 31656 MW; A7431C29117A503E CRC64;

Query Match 30.0%; Score 480; DB 1; Length 292;
Best Local Similarity 40.7%; Pred. No. 1.8e-30;
Matches 112; Conservative 50; Mismatches 87; Indels 26; Gaps 9;

QY 1 AQLSEFNFTKFDLQKDLIFOGDA-TSTNNVLQTLKDSGGNPGVSGRVLFSAPFHLW 59
| | | | | : | : | | | : | : | | | : | : | | | : | : | | | : | : |
Db 38 AEILSFSPKPKASQEDLULQGDALVSKGELQLTTVEN-GVPIWNSTGRALYAPVHIW 96
| | | | | : | : | | | : | : | | | : | : | | | : | : | | | : | : |
QY 60 ENSMA-VSSFTNLTIQISTPHYYAAGFAFFLPHDTVIPPNWSGKFLGLYSNVFRNS 118
| | | | | : | : | | | : | : | | | | | | | | | | | | | | | | |
Db 97 DKSTGRVASFATSFVVKAPVASKSADGIAFFLAPPNNQI-QGPGGGHGLFHS----- 150
| | | | | : | : | | | : | : | | | | | | | | | | | | | | | | |
QY 119 PTSENQSGDVNTDSRVAVVEFDTPPNANIDPNYRHIGIDVNSIKSKETARWONGKTA 178
| | | | | : | : | | | : | : | | | | | | | | | | | | | | | | |
Db 151 -----SGYNSQYIIIVDFDTHINA-WDPNTRHIGIDVNSINSTKTVTWQNGEYA 201
| | | | | : | : | | | : | : | | | | | | | | | | | | | | | | |
QY 179 TARISYNSAKSKSTVTTFYPGMEV-VALSVDVLDLHAELPEWVRVGLSASTG---EEKOKN 234
| | | | | : | : | | | : | : | | | | | | | | | | | | | | | | |
Db 202 NVLISYQAATETLVSLTPSSQTSYIILSAVDLSILPEWVRVGFATATGLTTQVVEH 261
| | | | | : | : | | | : | : | | | | | | | | | | | | | | | | |
QY 235 TTIWSFTSSLKNNVEKPEKEDMYIANYVRSYTWI 269
| | | | | : | : | | | : | : | | | | | | | | | | | | | | | | |
Db 262 DVLWSFTSTLTGTGCGAKDDNVHLV-----SYAFI 292
| | | | | : | : | | | : | : | | | | | | | | | | | | | | | | |

RESULT 15
LECS_MEDTR STANDARD; PRT; 280 AA.
AC Q01807;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Truncated lectin 2 precursor.
GN LEC2.
OS Medicago truncatula (Barrel medic).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales; Medicago.
OX NCBI_TaxID=3880;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Jemalong;
RX MEDLINE=92379255; PubMed=1511126;

```

```

RA Bauchowitz M.A., Barker D.G., Nadaud I., Rouge P., Lescure B.;
RT "Lectin genes from the legume Medicago truncatula.";
RL Plant Mol. Biol. 19:1011-1017(1992).
CC -1- MISCELLANEOUS: LEC2 IS PROBABLY NON FUNCTIONAL, SINCE A FRAMESHIFT
CC MUTATION LEADS TO PREMATURE TRANSLATION TERMINATION AFTER ONLY 98
CC AA. THE SEQUENCE BELOW IGNORES THIS FRAMESHIFT MUTATION.
CC -1- MISCELLANEOUS: BINDS ONE MANGANESE (OR OTHER TRANSITION METAL) ION
CC AND ONE CALCIUM ION. THE METAL IONS ARE ESSENTIAL FOR THE
CC SACCHARIDE-BINDING AND CELL-AGGLUTINATING ACTIVITIES.
CC -1- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60387; CAA42938.1; -
DR HSSP; P04122; lLOE.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
KW Lectin; Manganese; Calcium; Signal; Multigene family; Glycoprotein.
FT SIGNAL 1 26
FT CHAIN 27 280 TRUNCATED LECTIN 2.
FT METAL 148 148 MANGANESE (BY SIMILARITY).
FT METAL 150 150 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 152 152 CALCIUM (BY SIMILARITY).
FT METAL 154 154 MANGANESE (BY SIMILARITY).
FT METAL 158 158 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 170 170 MANGANESE (BY SIMILARITY).
FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 280 AA; 30473 MW; DB68690AD8015E81 CRC64;

Query Match 29.9%; Score 478; DB 1; Length 280;
Best Local Similarity 42.7%; Pred. No. 2.4e-30;
Matches 112; Conservative 37; Mismatches 79; Indels 34; Gaps 7;

QY 1 AQSLSFNFTRFDLDKDLIFOGDA-TSTNNVLQLTKLDSGGNPVCASVGRVLFSAFHLW 59
DB 27 AETTSFSTIKFVDPQKNLIFOGDKASTGKLEJSK-----AVKNSIGRALYSAPIHIW 80
QY 60 ENSM-AVSSFTNLTQISTPHYPYAADGAFFLAPHDTVIPPNWSGKFLGLY-SNVFRN 117
DB 81 DSKTGSVANFTTFTTITAPNTYNVADGLAFFIADTPDKPSIHGGYLGVDSDSKTYKK 140
QY 118 SPTSENSFGDVNTDSRVVAVFEFTFPNANIDPN-----YRHGIDVNSIKSKETARW 170
DB 141 S-----IQTVAVEIDTFYNAQWDPNPGNISSTGRHIGIDVNSIKSISTVPW 186
QY 171 EWONGKTATARIYSNASKSKTFTTFYFGMEVVALSHDVLHAELEPWVRVGLSASTGEE 230
DB 187 SLENNKANVAIGNGATNVLSDVDEYPLRIHYTLSHVVPKLVDPVPEWVRIGFSSTGAE 246
QY 231 KQKNTIISWFSFTSSL-----KNN 248
DB 247 YSAHDILSWFSFDKLNGLGFENN 268

```

Search completed: February 26, 2003, 16:52:56  
 Job time : 9.70037 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Database : PIR_73:**
1: pir1:**
2: pir2:**
3: pir3:**
4: pir4:**
```

## SUMMARIES

Result No.	Query Match	Score	%			DB	ID	Description
			Match	Length				
1	658	41.2	290	2	A34139	concanavalin A pre		
2	656	41.0	290	1	CVJBP	concanavalin A pre		
3	620	38.8	290	2	S66357	mannose/glucose-bi		
4	604.5	37.8	293	2	S66356	mannose/glucose-bi		
5	501.5	31.4	290	2	S66355	lectin-related sto		
6	500	31.3	287	2	JC5444	leukoagglutinin pr		
7	495	31.0	286	2	S48033	lectin precursor -		
8	489.5	30.6	285	2	S62690	agglutinin I precu		
9	485	30.3	247	2	JC2288	hemagglutinin - Ma		
10	483.5	30.2	285	2	S62691	agglutinin II prec		
11	481	30.1	233	1	LNLWBA	lectin precursor (		
12	477	29.8	279	2	T09620	probable lectin 2		
13	471	29.5	275	1	LNPM	lectin precursor -		
14	467.5	29.2	244	2	S23099	lectin I, anti-H(O		
15	459.5	28.7	249	2	J01981	lectin II - Scotch		
16	459	28.7	273	2	A57501	phytohemagglutinin		
17	452	28.3	233	1	FVFFBA	favin precursor -		
18	444.5	27.8	276	2	S1831	phytohemagglutinin		
19	439.5	27.5	272	2	B22826	phytohemagglutinin		
20	439	27.5	250	2	S16964	lectin - Scotch la		
21	436.5	27.3	249	2	JX0163	lectin II - furze		
22	435.5	27.2	285	2	S27365	lectin precursor -		
23	430	26.9	274	2	S1832	phytohemagglutinin		
24	427.5	26.7	275	2	A22826	phytohemagglutinin		
25	424.5	26.5	253	2	A53416	lectin II - lima b		
26	421.5	26.4	277	2	S25296	lectin LEC1 - barr		
27	419	26.2	275	2	A31972	lectin DB58 precu		
28	414	25.9	275	2	A29572	lectin precursor -		
29	413	25.8	236	1	LNOJ	lectin - common sa		

A: Reference number: A60636

A:Accession: A60636  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-290 <YA2>  
A:Experimental source: seed  
A:Note: the source was designated as Japanese jack bean  
A:C:Comment: Concanavalin A is formed from its precursor by post-translational cleavage  
CVJBWP and PIR:CVJB.

C:Superfamily: plant lectin  
C:Keywords: calcium; glycoprotein; homotetramer; lectin; manganese  
F:F.1-29/Domain: signal sequence #status predicted <SIG>  
F:F.148-149/Cleavage site: Asn-Val (unidentified proteinase) #status predicted  
F:F.152/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:F.163-164/Cleavage site: Asn-Ala (unidentified proteinase) #status predicted  
F:F.181-282/Cleavage site: Asn-Glu (unidentified proteinase) #status predicted

Query Match	41.2%;	Score 658;	DB 2;	Length 290;
Best Local Similarity	52.2%;	Pred. No. 8.5e-44;		
Matches 140;	Conservative 42;	Mismatches 66;	Indels 20;	Gaps 9;

[illegible]

## RESULT 1

A34139  
concanavalin A precursor - sword bean  
CCSpecies: Canavalia gladiata (sword bean)  
CCDate: 30-Jun-1991 #sequence\_revision 30-  
CCAccession: A34139; #G06036; JQ2130  
R:RYamauchi, D.; Minamikawa, T.  
FFBS Lett. 260, 127-130, 1990  
A:Title: Structure of the gene encoding con-  
A:Reference number: A34139; MUID:90127395;  
A:Accession: A34139  
A:Molecule type: DNA  
A:Residues: 1-260 <VAM>

A:Cross-references: EMBL:X16041; NID:g18009; PIDN:CAA34163.1; PID:g18010  
A:Note: the authors translated the codon ACG for residue 15 as Ile and GGA for residue  
R:Yamauchi, D.; Nakamura, K.; Asahi, T.; Minamikawa, T.  
Plant Cell Physiol. 30: 147-150. 1989

A:Reference number: A60636  
A:Accession: A60636  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-290 <YA>  
A:Experimental source: seed  
A:Note: the source was designated as Japanese jack bean  
C:Comment: Concanavalin A is formed from its precursor by post-translational cleavage  
CVJBP and PIR:CVJB.  
C:Superfamily: plant lectin  
C:Keywords: calcium; glycoprotein; homotetramer; lectin; manganese  
F:F:1-29/Domain: signal sequence #status predicted <SIC>  
F:F:148-149/Cleavage site: Asn-Val (unidentified proteinase) #status predicted  
F:F:152/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:F:163-164/Cleavage site: Asn-Ala (unidentified proteinase) #status predicted  
F:F:163-164/Cleavage site: Asn-Glu (unidentified proteinase) #status predicted  
F:F:281-282/Cleavage site: Asn-Glu (unidentified proteinase) #status predicted

Query Match	41.2%;	Score 658;	DB 2;	Length 290;
Best Local Similarity	52.2%;	Pred. No. 8.5e-44;		
Matches 140;	Conservative 42;	Mismatches 66;	Indels 20;	Gaps 9;

[illegible]

```

Db 209 KVTGTAHIYNSVGRKRLSNAVSYNGDSATVSYDVLNDVLPWVRVGLSASTGLYKETNT 268
QY 236 IISWTSSTSLKKNNEVKPEKEDMYIANVV 263
Db 269 ILSWSFTSKLKSNEIPD-----IATVV 290

RESULT 2
CWJB
concanavalin A precursor - jack bean
C:Species: Canavalia ensiformis (jack bean)
C:Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 28-May-1999
C:Accession: A03357; A60780; A60848
R:Carlington, D.M.; Auffret, A.; Hanke, D.E.
Nature 313, 64-67, 1985
A:Title: Polypeptide ligation occurs during post-translational modification of concanavalin A
A:Reference number: A03357; MUID:85086270; PMID:3965973
A:Accession: A03357
A:Molecule type: mRNA
A:Residues: 1-290 <CAR>
A:Cross-references: GB:X01632; NID:g312382; PIDN:CAA25787.1; PID:g17979
R:Chrispeels, M.J.; Hartl, P.M.; Sturm, A.; Faye, L.
J. Biol. Chem. 261, 10021-10024, 1986
A:Title: Characterization of the endoplasmic reticulum-associated precursor of concanavalin A
A:Reference number: A60780; MUID:86278043; PMID:3733700
A:Accession: A60780
A:Molecule type: protein
A:Residues: 30-41;153-169 <CHR>
R:Bowles, D.J.; Marcus, S.E.; Pappin, D.J.C.; Findlay, J.B.C.; Eliopoulos, E.; Maycox, H.
J. Cell Biol. 102, 1284-1297, 1986
A:Title: Posttranslational processing of concanavalin A precursors in jackbean cotyledon
A:Reference number: A60848; MUID:86168475; PMID:3958046
A:Accession: A60848
A:Molecule type: protein
A:Residues: 30-41;160-173 <BOW>
C:Comment: The mature chain (see PIR:CVJB) consists of residues 164-281 followed by 30-173
the signal sequence; cleavage after asparagines at positions 148, 163, and 281 is followed by
C:Superfamily: plant lectin
C:Keywords: glycoprotein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:149-163/Domain: glycopeptide #status predicted <GLP>
F:148-149/Cleavage site: Asn-Val (unidentified proteinase) #status experimental
F:152/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:163-164/Cleavage site: Asn-Ala (unidentified proteinase) #status experimental
F:281-282/Cleavage site: Asn-Glu (unidentified proteinase) #status experimental

Query Match 41.0%; Score 656; DB 1; Length 290;
Best Local Similarity 52.2%; Pred. No. 1.2e-43;
Matches 140; Conservative 41; Mismatches 67; Indels 20; Gaps 9;

QY 3 SLISNFTKFDLDKDLIFOGDATS--TNNVLQTLKLDGSGNPVGCASVGRVLFSAFFHLWEN 61
Db 36 ALHFENFQSKDKDLILQGDATTGTEGNLRTVRSSNGSPQGSVGRALFAPVHIVES 95
QY 62 SMWSSFTNLTITQISTP--HPYYAAGDGAFFLAPHDVTIPPNSWGKFLGLY--SNVERN 117
Db 96 SAVVASFEATFTFLKSPDSHP--ADGIAFFISNIDSIIPSGSGRLGLFPDANVIRN 152
QY 118 SPTSE-NGSGDVNTDSRVAAVEFTFPNANI--DPNRYRHIGIDVNSIKSKETARWEQNG 175
Db 153 STTIDFNAAY---NADT-IVAVELDTPTNTDIDGFSYPHIGIDIKSVRSKKTAKNNQNG 208
QY 176 KTATARIISYNSAKSKSTVTTFPGVGEVALSHDVLHAELPEWVRVGLSASTGEKQKNT 235
Db 209 KVGTAHIYNSVDKRLSAVSYNPADSATVSYDVLNDVLPWVRVGLSASTGLYKETNT 268
QY 236 IISWTSSTSLKKNNEVKPEKEDMYIANVV 263
Db 269 ILSWSFTSKLKSNEIPD-----IATVV 290

RESULT 3
mannose/glucose-binding lectin CLAI precursor - Cladrastis lutea
C:Species: Cladrastis lutea
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Aug-1999
C:Accession: S66357; S66301; S72502
R:van Damme, E.J.M.; Barre, A.; Bemer, V.; Rouge, P.; van Leuven, F.; Peumans, W.J.
Plant Mol. Biol. 29, 579-598, 1995
A:Title: A lectin and a lectin-related protein are the two most prominent proteins in
A:Reference number: S66299; MUID:96123235; PMID:8534854
A:Accession: S66357
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-290 <VAN>
A:Cross-references: EMBL:U21959
A:Experimental source: bark
A:Accession: S66301
A:Molecule type: protein
A:Residues: 36-46,'S',48,'NEA',52-54 <VAF>
R:van Damme, E.J.M.
submitted to the EMBL Data Library, March 1995
A:Reference number: S72502
A:Accession: S72502
A:Molecule type: mRNA
A:Residues: 1-153,'L',155-230 <VAV>
A:Cross-references: EMBL:U21959; NID:g1141758; PIDN:AAC49137.1; PID:g1141759
A:Experimental source: bark
C:Superfamily: plant lectin
C:Keywords: glycoprotein; lectin
F:1-35/Domain: signal sequence #status predicted <SIG>
F:36-290/Product: mannose/glucose-binding lectin CLAI #status experimental <MAT>

Query Match 38.8%; Score 620; DB 2; Length 290;
Best Local Similarity 49.8%; Pred. No. 7.7e-41;
Matches 136; Conservative 39; Mismatches 70; Indels 28; Gaps 8;

QY 1 AQLSUFNTKFDLDKDLIFOGDA--TSTNNVLQTLKLDGSGNPVGCASVGRVLFSAFFHL 58
Db 36 SDSLSFTFDNRPQDRILQGDAKISSGSLQUTKTDTSGKPVRSVGRALYITPLHL 95
QY 59 WENSM-AVSSFETNLTITQISTP--HPYYAAGDGAFFLAPHDVTIPPNSWGKFLGLYS--N 113
Db 96 WDSSTNRLASFQTTFTFLVSSPTNNP--GDGIAFFIAPPETTIPGSSGGLGLFSPDN 152
QY 114 VFRNSPTSENGSGDVNTDSRVAAVEFTFPNANIDPNYRHIGIDVNSIKSKETARWEQ 173
Db 153 AVNNSL-----NQIVAVEFTFVNNNDPSHRHIGIDVNTIKSSATVWQRE 199
QY 174 NGKTATARIISYNSAKSKSTVTTFYPM--EVALSHDVLHAELPEWVRVGLSASTGEE 230
Db 200 NGLSATQAISYNSDTKKLSVVSSYPNTQANEDYTVSYDVLKTELPEWVRVGFSGSTGY 259
QY 231 KQKNTIISWTSSTSLKKNNEVKPEKEDMYIANVV 263
Db 260 VQNHILSWTFNSNLQSSRAK--KEDIYIKRYV 290

RESULT 4
S66356
mannose/glucose-binding lectin CLAI precursor - Cladrastis lutea
C:Species: Cladrastis lutea
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Aug-1999
C:Accession: S66356; S66300
R:van Damme, E.J.M.; Barre, A.; Bemer, V.; Rouge, P.; van Leuven, F.; Peumans, W.J.
Plant Mol. Biol. 29, 579-598, 1995
A:Title: A lectin and a lectin-related protein are the two most prominent proteins in
A:Reference number: S66299; MUID:96123235; PMID:8534854
A:Accession: S66356
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-293 <VAN>
A:Cross-references: EMBL:U21958; NID:g1141756; PIDN:AAC49136.1; PID:g1141757
A:Experimental source: bark
A:Accession: S66300

```



A:Cross-references: EMBL:D17757; NID:g538528; PIDN:BAA04604.1; PID:g538529  
 A:Experimental source: inner bark  
 R:Tazaki, K.; Yoshida, K.; Shinohara, K.; Koshiba, T.; Yamamoto, N.  
 FEBS Lett. 377, 54-58, 1995  
 A:Title: Expression of cDNA for a bark lectin of Robinia in transgenic tobacco plants.  
 A:Reference number: S68376; MUID:96130320; PMID:8543018  
 A:Accession: S68376

A:Molecule type: protein  
 A:Residues: 32-40 <TAZ>  
 A:Experimental source: inner bark  
 C:Superfamily: plant lectin  
 C:Keywords: glycoprotein; lectin  
 F:1-31/Domain: signal sequence #status predicted <SIG>  
 F:32-286/Product: lectin precursor #status experimental <MAT>

Query Match 31.0%; Score 495; DB 2; Length 286;  
 Best Local Similarity 45.2%; Pred. No. 4e-31;  
 Matches 117; Conservative 39; Mismatches 81; Indels 22; Gaps 9;

Qy 3 SLSEFNFTKDLQKLIQFGDATTN-NVLQTLKDSGGNPVGASVGRVLSAPFHLWEN 61  
 ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
 Db 34 SLSEFPKFKHSQDLIFQSDALVTSKGVQLTVDN-GRPVYDSIGRVLYAAPFQIWD 92  
 Qy 62 SMA-VSSEFNLTQISTPHPYAAGDAFFLAPHDVTVIPNSWGKFLGLYSNVFRNSPT 120  
 :  
 Db 93 TTGNVASFVTSFIIKAPNEGKTADGLVFLAPVGS-TQPLKGGGLGLF----- 142  
 Qy 121 SENQSGDVNTDSRVVAVEFDTPFNANIDPNYRHIGIDVNSIKSKETARWQNGKTATA 180  
 :  
 Db 143 -KDESY---NKSQIVAVEFDTPFNANIDPNYRHIGIDVNSIKSKETARWQNGKTATA 198  
 :  
 Qy 181 RISYNSASKKSTVTTFYPGME-VVYALSHDYLHAELPEWVRVGLSASTG---EKQKNTI 236  
 ||||| :  
 Db 199 FISYEASTKSLTASLVPSLEKSFILSAIVDLKKVLPWVRVGVGTATTGLSEDIYVQNDV 258  
 Qy 237 ISWFTSSLK-NNEVKEPK 254  
 :  
 Db 259 LSWSFESNLPGGNSVASVK 277

RESULT 8  
 S62690

agglutinin I precursor - black locust  
 N:Alternate names: lectin  
 C:Species: Robinia pseudoacacia (black locust)  
 C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 07-May-1999  
 C:Accession: S62685  
 R:van Damme, E.J.M.; Barre, A.; Rouge, P.; van Leuven, F.; Peumans, W.J.  
 Plant Mol. Biol. 29, 1197-1210, 1995  
 A:Title: The seed lectins of black locust (robinia pseudoacacia) are encoded by two genes  
 A:Reference number: S62685; MUID:96191285; PMID:8616218  
 A:Accession: S62690  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-285 <VAN>  
 A:Cross-references: EMBL:U24249  
 A:Experimental source: seed  
 A:Accession: S62685  
 A:Molecule type: protein  
 A:Residues: 32-51 <VAN>  
 C:Superfamily: plant lectin  
 C:Keywords: glycoprotein; homotetramer; lectin  
 F:1-31/Domain: signal sequence #status predicted <SIG>  
 F:32-285/Product: agglutinin I #status experimental <MAT>

Query Match 30.6%; Score 489.5; DB 2; Length 285;  
 Best Local Similarity 43.6%; Pred. No. 1.1e-30;  
 Matches 113; Conservative 37; Mismatches 86; Indels 23; Gaps 8;

Qy 3 SLSEFNFTKDLQKLIQFGDATTN-NVLQTLKDSGGNPVGASVGRVLSAPFHLWEN 61  
 ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
 Db 34 SLSEFPKFKAPQYLIQFQDALVTSGVQLQINV-VNGVPPRSIGRALYAAPFQIWDN 92

Qy 62 SMA-VSSEFNLTQISTPHPYAAGDAFFLAPHDVTVIPNSWGKFLGLYSNVFRNSPT 120  
 :  
 Db 93 TTGNVASFVTSFIIKAPNEGKTADGLVFLAPVGS-TQPLKGGMLGIFKD----- 143  
 Qy 121 SENQSGDVNTDSRVVAVEFDTPFNANIDPNYRHIGIDVNSIKSKETARWQNGKTATA 180  
 :  
 Db 144 -----GSYNSKSNQIVAVEFDTPFNANIDPNYRHIGIDVNSIKSKETARWQNGKTATA 197  
 :  
 Qy 181 RISYNSASKKSTVTTFYPGMEVVVALSHD-VLHAELPEWVRVGLSASTGEEK---QKNTI 236  
 ||||| :  
 Db 198 FISYEASTKSLNASLVPSLETSTFIIHAIVDKVLDVLPWVRVGFSAATTGIDTGYVQNDV 257  
 :  
 Qy 237 ISWFTSSLK-NNEVKEPK 254  
 :  
 Db 258 LSWSFESNLPGGNSVASVK 276

RESULT 9  
 JC2268

hemagglutinin - Maackia amurensis  
 C:Species: Maackia amurensis  
 C:Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 07-May-1999  
 C:Accession: JC2268  
 R:Konami, Y.; Ishida, C.; Yamamoto, K.; Osawa, T.; Irimura, T.  
 J. Biochem. 115, 767-777, 1994  
 A:Title: A unique amino acid sequence involved in the putative carbohydrate-binding d  
 urensis hemagglutinin (MAH).  
 A:Reference number: JC2268; MUID:94375425; PMID:8089095  
 A:Accession: JC2268  
 A:Molecule type: protein  
 A:Residues: 1-247 <KON>  
 A:Experimental source: seed  
 C:Superfamily: plant lectin  
 C:Keywords: calcium binding; glycoprotein; manganese  
 F:11,177,189/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:125,127,138,143/Binding site: manganese (Glu, Asp, Asp, His) #status predicted  
 F:127,135,138/Binding site: calcium (Asp) #status predicted

Query Match 30.3%; Score 485; DB 2; Length 247;  
 Best Local Similarity 45.3%; Pred. No. 2e-30;  
 Matches 117; Conservative 39; Mismatches 74; Indels 28; Gaps 10;

Qy 1 AQSLSFNFTKDLQKLIQFGDATTN-NVLQTLKDSGGNPVGASVGRVLSAPFHLW 59  
 :  
 Db 1 SDELSFTINNFMPNQGDLDFQGVATVSPVTGVLQTS-EENGQPLEYSVGRVLTAPVRIW 59  
 Qy 60 ENSM-AVSSEFNLTQISTPHPYAAGDAFFLAPHDVTVIPNSWGKFLGLYSNVFRNS 118  
 :  
 Db 60 DSTTGAVASFSTFTVVKAAAG--ASDGLAFLLAPDPSQIFSGSVSKYLGLENN--SNS 115  
 Qy 119 PTSENSQSGDVNTDSRVVAVEFDTPFNANI---DPNRYHIGIDVNSIKSKETARWQNG 175  
 :  
 Db 116 DSS-----NQIVAVEFDTFGHSDPDWPNRYRHIGIDVNGIESIKTVQWDWING 164  
 Qy 176 KTATARIYSNSASKKSTVTTFYPGME---VVALSHDVLHAELPEWVRVGLSASTGEEKQ 232  
 :  
 Db 165 GVAFATITLAPNKTILASLVPSNQTSFIVAAS--VDLKEILPEWVRVGFSAATGAPKA 222  
 Qy 233 KNT--IISWSFTSSLKNN 248  
 :  
 Db 223 VETHDVRSWSFTSTLEAN 240

RESULT 10  
 S62691

agglutinin II precursor - black locust  
 N:Alternate names: lectin  
 C:Species: Robinia pseudoacacia (black locust)  
 C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 07-May-1999  
 C:Accession: S62691; S62686  
 R:van Damme, E.J.M.; Barre, A.; Rouge, P.; van Leuven, F.; Peumans, W.J.  
 Plant Mol. Biol. 29, 1197-1210, 1995  
 A:Title: The seed lectins of black locust (robinia pseudoacacia) are encoded by two g





Query Match 29.8%; Score 477; DB 2; Length 279;  
Best Local Similarity 43.5%; Pred. No. 9.7e-30;  
Matches 111; Conservative 36; Mismatches 78; Indels 30; Gaps 6;

QY 1 AQSLSFNFTFDLQDKLLIFOGDA-TSTNNVLOLTKLDSGGNPVCASGVRYLFSAPPHLW 59  
DB 27 AETTSFSITKFPQKNLIFOGDAETTSKGKLTGK-----AVKNSIGRALISAPHIW 80  
QY 60 ENSM-AVSSFETNLTQISTPHYYAAGDAFFLAPHDTVIPPNSWKGFLGLY-SNVFRN 117  
DB 81 DSKTGSVANPETTFETTTITAPNTYNVADGLAFFIAPIDTQPKNSQGGYLGVEDSKTYQE 140  
QY 118 SPTSENOFGDVNTDSRWAVEPDTFPNANIDP-----NYRHIGIDVNSIKKETARW 170  
DB 141 S-----IQTVAVEIDFTYVNDMDPKPGHNSISTGRHIGINVNSIKSITTPW 186  
QY 171 EWONGKTATARISSNASKSTVTFTFFPGMEVVALSHDVDLHAELPEWVRVGLSASTGEE 230  
DB 187 SLVNNKANVVGNGATNVLSDVEYPLVRHYHTLSHVPLKDVVPEWVRIGFSAATGDE 246  
QY 231 KQNTIISWSFTSSL 245  
DB 247 YAEHDFSWSFDSKL 261

RESULT 13  
LNPM  
lectin precursor - garden pea  
C:Species: Pisum sativum (garden pea)  
C:date: 03-Aug-1984 #sequence\_revision 03-Aug-1984 #text\_change 16-Jul-1999  
C:Accession: A26844; A92438; A90627; S22074; S42645; A03359  
R:Gatehouse, J.A.; Bown, D.; Evans, I.M.; Gatehouse, L.N.; Jobes, D.; Preston, P.; Croy,  
Nucleic Acids Res. 15, 7642, 1987  
A:title: Sequence of the seed lectin gene from pea (Pisum sativum L.).  
A:Reference number: A26844; MUID:88015625; PMID:3658708  
A:Accession: A26844  
A:Molecule type: DNA  
A:Residues: 1-275 <GAT>  
A:Cross-references: GB:Y00440; NID:g20769; PIDN:CAA68497.1; PID:g20770  
R:Higgins, T.J.V.; Chandler, P.M.; Zurawski, G.; Button, S.C.; Spencer, D.  
J. Biol. Chem. 258, 9544-9549, 1983  
A:title: The biosynthesis and primary structure of pea seed lectin.  
A:Reference number: A92438; MUID:83285760; PMID:6688253  
A:Accession: A92438  
A:Molecule type: mRNA  
A:Residues: 1-275 <HTG>  
A:Experimental source: cv. p1/G 086  
A:Note: the authors translated the codon CAA for residues 5 and 7 as Glu  
R:Richardson, C.; Behnke, W.D.; Freisheim, J.H.; Blumenthal, K.M.  
Biochim. Biophys. Acta 537, 310-319, 1978  
A:title: The complete amino acid sequence of the alpha-subunit of pea lectin, Pisum sati  
A:Reference number: A90627; MUID:79082912; PMID:728447  
A:Accession: A90627  
A:Molecule type: protein  
A:Residues: 218-268,270-271 <RIC>  
R:de Pater, B.S.; Pham, K.T.; Katagiri, F.; Chua, N.H.; Kijne, J.W.  
submitted to the EMBL data Library, May 1992  
A:description: Seed-specific and developmental regulated expression.  
A:Reference number: S22074  
A:Accession: S22074  
A:Molecule type: DNA  
A:Residues: 1-275 <DEP>  
A:Cross-references: EMBL:X56358; NID:g562782; PIDN:CAA47011.1; PID:g20804  
A:Experimental source: var. Feldham First  
R:Roedemaeker, F.J.; Richardson, M.; Diaz, C.L.; de Pater, B.S.; Kijne, J.W.  
Plant Mol. Biol. 24, 75-81, 1994  
A:title: Pea (Pisum sativum L.) seed isolectins 1 and 2 and pea root lectin result from  
A:Reference number: S42645; MUID:94154245; PMID:8111028  
A:Accession: S42645  
A:Molecule type: mRNA  
A:Residues: 31-54;56-91;100-113;116-129;135-149;154-163;166-183;185-187;197-200;204-211  
C:Genetics:  
A:Gene: Leca

```
Db 62 TGRVASETFSFVVKDEPKSNGVDLTFLLAPANSQIPSGSSAGLFLGN-----S 114
QY 121 SENOSFGDVNTDSRVAVVEFTFPNANI---DPNVRHIGIDVNSIKSKETARWEWQNGKT 177
Db 115 SDNKS-----SNQIIAIEFFTYFGKTYNPWDPDFKHIGVDVNSIKSIKTIVKWDNRNGEV 168
QY 178 ATARISYNSASKSKSTVTTFYFG- --MEVVALSHDVLHAELEPEWVRVGLSASTGE--EKQ 232
Db 169 ANVITYRAPTKSLTSLSYSPSDQTSNIVTAS--VDLKAILPEWVSQVSGVGNAAEFE 226
QY 233 KNTIISWSTSLKNN 248
Db 227 THDVLSWYFTSLEAN 242
```

## RESULT 15

```
QJ1981
lectin II - Scotch broom
N:Alternate names: 2-acetamido-2-deoxy-D-galactose binding lectin II
C:Species: Cytisus scoparius (Scotch broom)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 07-May-1999
C:Accession: JQ1981
R:Konami, Y.; Yamamoto, K.; Osawa, T.; Irimura, T.
J. Biochem. 112, 366-375, 1992
A:Title: The primary structure of the Cytisus scoparius seed lectin and a carbohydrate-
A:Reference number: JQ1981; MUID:93054441; PMID:1429525
A:Accession: JQ1981
A:Molecule type: protein
A:Residues: 1-249 <KON>
A:Experimental source: seed
C:Superfamily: plant lectin
C:Keywords: lectin
```

```
Query Match 28.7%; Score 459.5; DB 2; Length 249;
Best Local Similarity 43.2%; Pred. No. 1.9e-28;
Matches 112; Conservative 43; Mismatches 75; Indels 29; Gaps 12;

QY 1 AQSLSFNFTKFDLQKDLIF-QGDATST-NNVLQTLKLDGNGNPVGASVGRVLFSAPEHL 58
Db 1 SEELSFSTKTKTDOKNLIILEQDALITPTGKLQLTIVEN-GKPAAYSIGRAIYSTPIHI 59
QY 59 WENMA-VSSPETNLTIOIS-TPHPYAA-DGFAFFLAPHDTVIPPNWSGKFLGLYSNVF 115
Db 60 WDKSTGDEASATFTSFVISDAPNPSTAAATDGLAFFLAPADT--QPQAGGYLGLF--- 113
QY 116 RNSPTSENOFGDVNTDSRVAVVEFTFPNANI-DPNR-HIGIDVNSIKSKETARWEWQN 174
Db 114 -----EKDSSY---NSSNQIVAVEFTTYNSAWDPQTNPHIGIDVNTIKSKVSSWGFKN 165
QY 175 GKTATARISYNSASKSKSTVTTFYPGMEV-----VALSHDVLHAELEPEWVRVGLSASTGE 229
Db 166 GNVATVLTITYOPSSKSLVASLVYPSQTSKTSYIISANVDLKATVPWVRIGFSATGQ 225
QY 230 EK---QKNTIISWSTSL 245
Db 226 TDNVIETHDILSWSPKSL 244
```

Search completed: February 26, 2003, 16:51:32  
Job time : 23.6966 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:40:30 : Search time 13.5131 Seconds  
(without alignments)  
574.823 Million cell updates/sec

Title: US-09-476-485A-2

Perfect score: 1370

Sequence: 1 AGSLSFSTKFDPNQEDLIF.....LWTVAKKENKXITRGVL 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1362	99.4	270	4	US-08-881-189B-2
2	1362	99.4	286	4	US-08-881-189B-23
3	575.5	42.0	132	4	US-08-881-189B-12
4	463	33.8	285	3	US-09-141-821-2
5	441	32.2	285	3	US-09-141-821-4
6	440	32.1	285	3	US-09-141-821-1
7	439	32.0	285	3	US-09-141-821-5
8	430.5	31.4	105	4	US-08-881-189B-13
9	420	30.7	262	2	US-08-038-761A-1
10	321.5	23.5	274	3	US-09-141-821-3
11	244	17.8	632	4	US-09-228-986-77
12	107.5	7.8	1095	4	US-09-206-942-45
13	107.5	7.8	1101	4	US-09-206-942-43
14	105	7.7	18	4	US-08-881-189B-14
15	101	7.4	304	1	US-07-851-976B-8
16	101	7.4	304	1	US-08-291-609-8
17	101	7.4	304	1	US-08-401-136-8
18	101	7.4	304	1	US-08-850-554-8
19	97	7.1	1876	2	US-08-609-049A-12
20	97	7.1	1876	2	US-08-609-049A-28
21	97	7.1	1876	2	US-09-170-996-12
22	97	7.1	1876	4	US-09-170-996-28
23	91	6.6	675	4	US-08-426-509A-4
24	91	6.6	675	5	PCT-US95-05008-4
25	91	6.6	775	1	US-07-603-133B-15
26	90	6.6	775	1	US-07-603-133B-16
27	89	6.5	724	3	US-09-121-964-1

28 89 6.5 860 1 US-08-092-817-4 Sequence 4, Appli  
29 89 6.5 860 4 US-08-485-128-4 Sequence 4, Appli  
30 89 6.5 860 4 US-09-804-778A-8 Sequence 8, Appli  
31 89 6.5 1222 4 US-09-206-942-37 Sequence 37, Appli  
32 89 6.5 1228 4 US-09-206-942-34 Sequence 34, Appli  
33 89 6.5 1410 2 US-08-470-058-4 Sequence 4, Appli  
34 89 6.5 1410 3 US-09-037-188-4 Sequence 4, Appli  
35 89 6.5 1410 4 US-09-285-310-4 Sequence 8, Appli  
36 86 6.3 790 3 US-08-537-361E-8 Sequence 4, Appli  
37 85.5 6.2 463 2 US-08-853-659A-52 Sequence 52, Appli  
38 84.5 6.2 616 4 US-09-298-367B-11 Sequence 11, Appli  
39 84.5 6.2 1250 1 US-08-441-139-9 Sequence 9, Appli  
40 84 6.1 790 4 US-08-817-707-8 Sequence 8, Appli  
41 84 6.1 1002 4 US-09-268-347-24 Sequence 24, Appli  
42 83 6.1 1104 4 US-09-268-347-28 Sequence 28, Appli  
43 83 6.1 1104 4 US-09-268-347-34 Sequence 34, Appli  
44 82.5 6.0 514 3 US-08-467-023-134 Sequence 134, App  
45 82.5 6.0 853 4 US-08-913-880C-17 Sequence 17, Appli

## ALIGNMENTS

RESULT 1  
US-08-881-189B-2  
; Sequence 2 Application US/08881189B  
; Patent No. 6310195  
; GENERAL INFORMATION:  
; APPLICANT: Colucci et al.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED  
; PROGENITOR CELL PRESERVATION FACTOR  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron, LLP  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11753  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/881,189B  
; FILING DATE: June 24, 1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: 381-44 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 270 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-881-189B-2

Query Match 99.4%; Score 1362; DB 4; Length 270;  
Best Local Similarity 99.6%; Pred. No. 4.9e-128;  
Matches 263; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGSLSFSTKFDPNQEDLIFQGHATSTNNVLQTKLDSAGNPVSSSAGRVLYSAPRLWE 60  
Db 1 AQSLSFSTKFDPNQEDLIFQGHATSTNNVLQTKLDSAGNPVSSSAGRVLYSAPRLWE 60

100

```

US-09-141-821-2
; Sequence 2, Application US/09141821
; Patent No. 6110891
; GENERAL INFORMATION:
; APPLICANT: Arpad Janos PUSZTAI
; APPLICANT: Szuzanna Magdalena BARDOZ
; APPLICANT: Richard Michael John PALMER
; APPLICANT: Neil William FISH
; APPLICANT: Gyorgy J. KOTELES
; TITLE OF INVENTION: LECTIN COMPOSITIONS AND USES THEREOF
; FILE REFERENCE: 48656
; CURRENT APPLICATION NUMBER: US/09/141,821
; CURRENT FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Human
US-09-141-821-2

Query Match 33.8%; Score 463; DB 3; Length 285;
Best Local Similarity 43.9%; Pred. No. 2.9e-36;
Matches 115; Conservative 40; Mismatches 81; Indels 26; Gaps 10;

QY 2 GSLSFSTKFDPNQEDLIFQGHATSTN-NVLQVTKLDSAGNPVSSAGRVLYSAPLRWE 60
D 33 GSLSFSPKPKHSPDLIFQSDALVTSKGVQLTVD--GRVYDSIGRVLYAAPFQIWD 90
QY 61 D-SAVLTSTFTIINFELSTPYTSRIADGLAFFIAPPDSVISYH-GGFGLFPNANTLNS 118
D 91 STTGNVASFVTSFIIKAPNECKTADGLVFLAPVGSQPLKGGGLGLF----- 141
QY 119 STSENOTTTKAASNNVAVFEDTLYNPDYDGNPIHIGIDVNSIRSKVTAKWDQNGKIA 178
D 142 ---KDESINK--SNQIVAVEEDTFERNVAV-DPNCIHRMGIDVNSIQSVRTVRWDWANGEVA 195
QY 179 TAHISYNSVKRLSVTSYYAG-SKPAFLSYDIELHTVLPENVRVGLSASTGQDK---ERN 234
D 196 NVFISYEASTKSLTASLVYPSLEKSFILSAIVDLKVLPEWVRVGFATTGLSEDIYQTN 255
QY 235 TVHWSFTSSL--WTNVAKKEN 254
D 256 DVLSWSFESNLPCCGNSVASVKN 277

RESULT 5
US-09-141-821-4
; Sequence 4, Application US/09141821
; Patent No. 6110891
; GENERAL INFORMATION:
; APPLICANT: Arpad Janos PUSZTAI
; APPLICANT: Szuzanna Magdalena BARDOZ
; APPLICANT: Richard Michael John PALMER
; APPLICANT: Neil William FISH
; APPLICANT: Gyorgy J. KOTELES
; TITLE OF INVENTION: LECTIN COMPOSITIONS AND USES THEREOF
; FILE REFERENCE: 48656
; CURRENT APPLICATION NUMBER: US/09/141,821
; CURRENT FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Human
US-09-141-821-4

Query Match 32.2%; Score 441; DB 3; Length 285;
Best Local Similarity 42.5%; Pred. No. 4.5e-36;
Matches 111; Conservative 42; Mismatches 84; Indels 24; Gaps 9;

QY 2 GSLSFSTKFDPNQEDLIFQGHATSTN-NVLQVTKLDSAGNPVSSAGRVLYSAPLRWE 60

```

```

D 33 GSLSFSEFPKFAFPQPYLIFQDALVTSTGVQLTNNVN-GVPPRRSIGRALLYAAPFQIWD 91
QY 61 D-SAVLTSTFTIINFELSTPYTSRIADGLAFFIAPPDSVISYHGGFLGFLFPNANTLNS 119
D 92 NITGNVASFVTSFIIQAPNPATTDGLAFFLAPVDTQPGDLGGLGIFKDG----- 144
QY 120 TSENOTTTKAASNNVAVFEDTLYNPDYDGNPIHIGIDVNSIRSKVTAKWDQNGKIA 179
D 145 -SYNK-----SNQIVAVEEDTFESNIHF-DPKGRHMGINVSIVSVKTPVNPWTNGEVAN 196
QY 180 AHISYNSVKRLSVTSYYAGSKPATLSYDI-ELHTVLPENVRVGLSASTGQDK---ERN 235
D 197 VFISYEASTKSLNASLVYPSLETSTFIIHAIVDVKDVLPEWVRVGFSAITGIDTGYVQTN 256
QY 236 VHSWSFTSSL--WTNVAKKEN 254
D 257 VLSWSFESNLPCCGNSVASVKN 277

RESULT 6
US-09-141-821-1
; Sequence 1, Application US/09141821
; Patent No. 6110891
; GENERAL INFORMATION:
; APPLICANT: Arpad Janos PUSZTAI
; APPLICANT: Szuzanna Magdalena BARDOZ
; APPLICANT: Richard Michael John PALMER
; APPLICANT: Neil William FISH
; APPLICANT: Gyorgy J. KOTELES
; TITLE OF INVENTION: LECTIN COMPOSITIONS AND USES THEREOF
; FILE REFERENCE: 48656
; CURRENT APPLICATION NUMBER: US/09/141,821
; CURRENT FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Human
US-09-141-821-1

Query Match 32.1%; Score 440; DB 3; Length 285;
Best Local Similarity 42.5%; Pred. No. 5.6e-36;
Matches 111; Conservative 40; Mismatches 86; Indels 24; Gaps 9;

QY 2 GSLSFSTKFDPNQEDLIFQGHATSTN-NVLQVTKLDSAGNPVSSAGRVLYSAPLRWE 60
D 33 GSLSFSEFPKFAFPQPYLIFQDALVTSTGVQLTNNVN-GVPSGKSLGRALLYAAPFQIWD 91
QY 61 D-SAVLTSTFTIINFELSTPYTSRIADGLAFFIAPPDSVISYHGGFLGFLFPNANTLNS 119
D 92 STTGNVASFVTSFIIQAPNPATTDGLAFFLAPVDTQDVGGLMGIFPKD-GYFNK-- 148
QY 120 TSENOTTTKAASNNVAVFEDTLYNPDYDGNPIHIGIDVNSIRSKVTAKWDQNGKIA 179
D 149 -----SNQIVAVEEDTFESNIHF-DPKGRHMGINVSIVSVKTPVNPWTNGEVAN 196
QY 180 AHISYNSVKRLSVTSYYAGSKPATLSYDI-ELHTVLPENVRVGLSASTGQDK---ERN 235
D 197 VFISYEASTKSLTASLVYPSLETSTFIIHAIVDVKDVLPEWVRVGFSAITGIDTGYVQTN 256
QY 236 VHSWSFTSSL--WTNVAKKEN 254
D 257 VLSWSFESNLPCCGNSVASVKN 277

RESULT 7
US-09-141-821-5
; Sequence 5, Application US/09141821
; Patent No. 6110891
; GENERAL INFORMATION:
; APPLICANT: Arpad Janos PUSZTAI

```





ORGANISM: Bauhinia purpurea  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:

US-08-038-761A-1

Query Match 30.78; Score 420; DB 2; Length 262;  
Best Local Similarity 39.08; Pred. No. 4.9e-34;  
Matches 97; Conservative 46; Mismatches 82; Indels 24; Gaps 6;  
QY 6 FSFTKF---DPNQEDLIFOGHATSTNNVQVTKLDSAGNPVSSSAGRVLYSAPLRLWED 61  
Db 8 FTTFNFSNQENGEIIFIGNAYTFCALRLTRIGEDGILKSNAGQASRPFVLWDS 67  
QY 62 SAVTSPDITNF---EISTPYTSRIADGLAFFIAPPDSVSYHGGFLGLFPNANTLNS 118  
Db 68 TGHVASEYTSFSFVRSIDVPHIT--ADGEAFFLAPVDSVKDYGGCLGLF----- 116  
QY 119 STSENQTTTKAASSNVAVEEDTYLNPDPYDGNPIHIGIDVNSIRSKVTAKDWQNGKI- 177  
Db 117 ---RYKTATPSKNQVAVVEEDTWPNTESDLRPHIGINVTSTVSATTRWDNDAYVT 173  
QY 178 -ATAHISYNSVKRLSVTSYVYAGSKPATLSYDIELHTVLPWVRVGLSASTGQDKERNTV 236  
Db 174 KSTAHITYDATSKILTYLLTYDNGRHHYQLSHVVDLPKILPVRIGFSGGTGFN-ETQVI 232  
QY 237 HSWSFSTSL 245  
Db 233 LSWSFSTSL 241

RESULT 10  
US-09-141-821-3  
Sequence 3, Application US/09141821  
Patent No. 6110891  
GENERAL INFORMATION:  
APPLICANT: Arpad Janos PUSZTAI  
APPLICANT: Szuzanna Magdalena BARDOZ  
APPLICANT: Richard Michael John PALMER  
APPLICANT: Neil William FISH  
APPLICANT: Gyorgy J. KOPELES  
TITLE OF INVENTION: LECTIN COMPOSITIONS AND USES THEREOF  
FILE REFERENCE: 48656  
CURRENT APPLICATION NUMBER: US/09/141.821

CURRENT FILING DATE: 1998-08-28  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 274  
TYPE: PRT  
ORGANISM: Human  
US-09-141-821-3

Query Match 23.58; Score 321.5; DB 3; Length 274;  
Best Local Similarity 36.88; Pred. No. 3.6e-24;  
Matches 93; Conservative 38; Mismatches 87; Indels 35; Gaps 11;  
QY 4 LSFSTKFDPNQEDLIFOGHAT-STNNVQVTKLDSAGNPVSSSAGRVLYSAPLRLWEDS 62  
Db 33 ISFNFTNFRGDQGVTLGQANIMANGILALT---NHTNP-TWNTGRALYSKRPVPIW-DS 87  
QY 63 AV--LTSFDTIINFESTPYTSRIADGLAFFIAP---PDSVSYHGGFLGLFPNANTLN 116  
Db 88 ATGNVASEFTVSFVVKKEIKGGIPADGIVFEFLAPEARIPDNSA---GGQLGIV-NANKAY 143  
QY 117 NSSTSENQTTTKAASSNVAVEEDTYLNPDPYDGNPIHIGIDVNSIRSKVTAKDWQNGK 176  
Db 144 NP-----FVGVFEFTYSN--NWDPKSAHIGIDASSLSLRTVKNWKNVSGS 186  
QY 177 IATAHISYNSVKRLSVTSYVYAGSKPATLSYDIELHTVLPWVRVGLSA---STGQDKERN 234  
Db 187 LVKVSIIYDSLSKTLSTLVVVTHENGQISTIAQVVDKAVLGEKVRVGTAAATTGRIYVELY 246  
QY 235 TVHSWSTSLWT 247  
Db 247 DIHWSFTSLVT 259

RESULT 11  
US-09-228-986-77  
Sequence 77, Application US/09228986  
Patent No. 6359198  
GENERAL INFORMATION:  
APPLICANT: Strabala, Timothy  
APPLICANT: Nieuwenhuizen, Niels  
TITLE OF INVENTION: Compositions Isolated from Plant Cells  
TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling  
FILE REFERENCE: 11000/1020  
CURRENT APPLICATION NUMBER: US/09/228, 986  
CURRENT FILING DATE: 1999-01-12  
NUMBER OF SEQ ID NOS: 130  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 77  
LENGTH: 632  
TYPE: PRT  
ORGANISM: Pinus radiata  
US-09-228-986-77

Query Match 17.88; Score 244; DB 4; Length 632;  
Best Local Similarity 28.48; Pred. No. 6.9e-16;  
Matches 73; Conservative 56; Mismatches 88; Indels 40; Gaps 10;  
QY 5 SFSFTKFDPNQEDLIFOGHAT--STNNVQVTKLDSAGNPVSSSAGRVLYSAPLRLWEDS 62  
Db 34 TFLFDGF--NGTNLILEANASVIGSESVLSLT-----NHSHEFMLGRALYAAPVQM-KNN 85  
QY 63 AVLTSPDITNFEISTPYTSRIADGLAFFIAPPDSVSYHGGFLGLFPNANTLNSSTS 121  
Db 86 HTVSPSFSTFTFVSIVPPPSNEGGLGLAFIMTPTSPMGAQPVQYLGLL----- 133  
QY 122 ENOTTTKAASSNVAVEEDTYLNPDPYDGNPIHIGIDVN---STRSKVTAKDWQ----- 173  
Db 134 -NLTSNGQPNHFLFAVEEDTIMNVEFKDPDRNHGVQDVINSLSVQETAGTWNGEERHEL 192  
QY 174 ---NGKIATAHISYNSVKRLSVTSYVYAG---SKPATLSYDIELHTVLPWVRVGLSAS 226  
Db 193 NLRSGRNIQAWIDYDHLLESSLNVTITVAGLPRPQRP-LISLQIDLQNLVIEKMLVGFSA 251

```
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-206-942-43

Query Match      7.8%; Score 107.5; DB 4; Length 1095;
Best Local Similarity 24.1%; Pred. No. 0.066; 99; Indels 53; Gaps 11;
Matches 60; Conservative 37; Mismatches 99; Indels 53; Gaps 11;

RESULT 12
US-09-206-942-45
; Sequence 45, Application US/09206942
; Patent No. 6432669
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High
; TITLE OF INVENTION: Molecular Weight Proteins
; FILE REFERENCE: 1038-861 MIS:jb
; CURRENT APPLICATION NUMBER: US/09/206,942
; CURRENT FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: 09/167,568
; EARLIER FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-206-942-45

Query Match      7.8%; Score 107.5; DB 4; Length 1095;
Best Local Similarity 24.1%; Pred. No. 0.066; 99; Indels 53; Gaps 11;
Matches 60; Conservative 37; Mismatches 99; Indels 53; Gaps 11;

; 26 STNNVLOVTKLDSAGNPVSSSAGRVLYSAPLR---LWEDSAVLTSFDTIINFEISTPYTS 82
; 230 STNSYLNIT-IDNSGRSPGAGPLVRRSLGNGISFNNDTVFNVASGSAVNFISIKPIVS 288
; 83 RTADGLAFFIAPDPSVISYHGGFLGLFPNANTLN-----NSSTSENQT---TTK 128
; 289 NVHDGNHTLFGNVSVLG--GGDVNFHFENASSNHWTHGVVKSQNFNASEGSSLRFKSE 346
; 129 AASNNVVAEFTYLPDYGDPNYIHI-GIDVNSIRSKVTAK-WDWQNGKIATAHISYNS 186
; 347 GSTRTAFIESDLTNATGNSISLNOVAGIDGNLQKSLVANKNITFEGGNITLA----- 400
; 187 VSKRLSVTSYYAGSKPATLSYDIELHTVLPWVRVGL-SASTGQDKERNTVHWSFTSSL 245
; 401 -----ADKKP-----IEIKGNITVKEGANVTLRSANYGNDKSALSIRG----- 438
; 246 WTNVAKKEN 254
; 439 --NVTNKGK 445

RESULT 13
US-09-206-942-43
; Sequence 43, Application US/09206942
; Patent No. 6432669
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High
; TITLE OF INVENTION: Molecular Weight Proteins
; FILE REFERENCE: 1038-861 MIS:jb
; CURRENT APPLICATION NUMBER: US/09/206,942
; CURRENT FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: 09/167,568
; EARLIER FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 1101
```

```
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-206-942-43

Query Match      7.8%; Score 107.5; DB 4; Length 1101;
Best Local Similarity 24.1%; Pred. No. 0.067; 99; Indels 53; Gaps 11;
Matches 60; Conservative 37; Mismatches 99; Indels 53; Gaps 11;

; 26 STNNVLOVTKLDSAGNPVSSSAGRVLYSAPLR---LWEDSAVLTSFDTIINFEISTPYTS 82
; 236 STNSYLNIT-IDNSGRSPGAGPLVRRSLGNGISFNNDTVFNVASGSAVNFISIKPIVS 294
; 83 RIADGLAFFIAPDPSVISYHGGFLGLFPNANTLN-----NSSTSENQT---TTK 128
; 295 NVHDGNHTLFGNVSVLG--GGDVNFHFENASSNHWTHGVVKSQNFNASEGSSLRFKSE 352
; 129 AASNNVVAEFTYLPDYGDPNYIHI-GIDVNSIRSKVTAK-WDWQNGKIATAHISYNS 186
; 353 GSTRTAFIESDLTNATGNSISLNOVAGIDGNLQKSLVANKNITFEGGNITLA----- 406
; 187 VSKRLSVTSYYAGSKPATLSYDIELHTVLPWVRVGL-SASTGQDKERNTVHWSFTSSL 245
; 407 -----ADKKP-----IEIKGNITVKEGANVTLRSANYGNDKSALSIRG----- 444
; 246 WTNVAKKEN 254
; 445 --NVTNKGK 451

RESULT 14
US-08-881-189B-14
; Sequence 14, Application US/08881189B
; Patent No. 6310195
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; TITLE OF INVENTION: PROGENITOR CELL PRESERVATION FACTOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,189B
; FILING DATE: June 24, 1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-881-189B-14

Query Match      7.7%; Score 105; DB 4; Length 18;
```

Best Local Similarity 100.0%; Pred. No. 0.00024;  
Matches 18; Conservative 0; Mismatches 0;

QY 142 YLNPDYGDPNYIHIGIDV 15  
| | | | | | | | | | | | | | | |  
Db 1 YLNPDYGDPNYIHIGIDV 18

QY 216 ---PEWVRV-GLSASTQDKERNTVHWSFTSSLW 246  
: | | | : | | : | | : : : |  
Db 241 ETCKQWSGVFGYDYSAPKTEANTPOT-NYETIWI 274

Search completed: February 26, 2003, 16:41:23  
Job time : 17.5131 secs

```

RESULT 15
US-07-851-976B-8
; Sequence 8, Application US/07851976B
; Patent No. 5426043
; GENERAL INFORMATION:
; APPLICANT: De Graaff, Leendert H.
; APPLICANT: Visser, Jacob
; APPLICANT: Van Den Broeck, Henriette C.
; APPLICANT: Strozyk, Francois
; APPLICANT: Kormelink, Felix J.M.
; APPLICANT: Boonman, Johannes C.P.
; TITLE OF INVENTION: CLONING AND EXPRESSION OF ACETYL XYLAN
; TITLE OF INVENTION: ESTERASES FROM FUNGAL ORIGIN
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
;

```

Query Match	7.4%	Score 101;	DB 1;	Length 304;
Best Local Similarity	21.1%;	Pred. No. 0.043;		
Matches	58;	Conservative 46;	Mismatches 95;	Indels 76; Gaps 15;
QY	23	HATSTNNVL-QVTKLDSAGNPVSSSAGRVLVSAPLRWEDSAVLTSPTDITNFEISTPYT	81	
Ddb		:    :    :    :    :    :		
QY	25	HAVKRSGLQQVT--DFGDNP--TNVGMIY-VPNNLASNPGIIVA-----IHCTGT---	72	
Ddb		:        :    :    :    :    :    :		
QY	82	SRIADGLAFIAPPDSVISYHGFGFLGFPA-----NTLNNSSTSENQ	124	
Ddb		:  :    :    :    :    :    :		
QY	73	-----GPGYISASPYATLSEQGFIIVIPSSPYGGCWVDSSOATLTHNGGNSIANM	127	
Ddb		:    :    :    :    :    :    :    :		
QY	125	TT-----TKAASNVAVEFDYLPNDGDPNYIHIGIDVNSIRSK	165	
Ddb		:    :    :    :    :    :    :    :		
QY	128	VTWITISEYGADSSKFVFVTGSSSCAMLTNWAA---TY--PELEAAATVYSGVAGCFYSN	182	
Ddb		:    :    :    :    :    :    :    :		
QY	166	VTAKDWQN----GKTATAHISYNSVSKRLSVTSYYAGSKPATLSYDIELHTVL-----	215	
Ddb		:    :    :    :    :    :    :    :		
QY	183	TNOVDGNWSTCAQGDIVITPEHWASTAE--AMYSGYSGSRPRMQIYHGTLLHTLYPONYY	240	
Ddb		:    :    :    :    :    :    :    :		



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:40:33 ; Search time 9.22847 Seconds  
(without alignments)  
1079.114 Million cell updates/sec

Title: US-09-476-485A-2

Perfect score: 1370

Sequence: 1 AGSLSFSTKFPDQEDLIF.....LWTNVAKKENKYYITRGVL 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174566 seqs, 37721826 residues

Total number of hits satisfying chosen parameters: 174566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PTCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PTCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	208.5	15.2	669	10 US-09-862-027-27	Sequence 27, Appl
2	103.5	7.6	639	10 US-09-782-906-2	Sequence 2, Appl
3	100.5	7.3	2344	10 US-09-815-242-12713	Sequence 12713, A
4	99.5	7.3	639	10 US-09-782-906-3	Sequence 3, Appl
5	99.5	7.3	639	10 US-09-782-906-4	Sequence 4, Appl
6	99.5	7.3	639	10 US-09-782-906-5	Sequence 5, Appl
7	92	6.7	415	9 US-09-893-519A-41	Sequence 41, Appl
8	91.5	6.7	1723	10 US-09-841-132-394	Sequence 394, Appl
9	91	6.6	433	10 US-09-801-368-226	Sequence 226, Appl
10	91	6.6	675	9 US-10-186-399-3	Sequence 3, Appl
11	91	6.6	675	9 US-09-977-260-4	Sequence 4, Appl
12	91	6.6	675	10 US-09-977-269-4	Sequence 4, Appl
13	90.5	6.6	1723	10 US-09-841-132-395	Sequence 395, Appl
14	89.5	6.5	375	10 US-09-886-468-20	Sequence 20, Appl
15	89.5	6.5	632	10 US-09-853-533A-8	Sequence 8, Appl
16	89	6.5	860	10 US-09-824-637-4	Sequence 4, Appl
17	88	6.4	596	9 US-10-063-547-100	Sequence 100, Appl
18	88	6.4	596	9 US-10-174-590-310	Sequence 310, Appl
19	88	6.4	596	9 US-10-176-758-310	Sequence 310, Appl

20	88	6.4	596	9 US-10-063-616-100	Sequence 100, Appl
21	88	6.4	596	9 US-10-175-737-310	Sequence 310, Appl
22	88	6.4	596	9 US-10-063-502-100	Sequence 100, Appl
23	88	6.4	596	9 US-10-173-706-310	Sequence 310, Appl
24	88	6.4	596	9 US-10-175-738-310	Sequence 310, Appl
25	88	6.4	596	9 US-10-175-752-310	Sequence 310, Appl
26	88	6.4	596	9 US-10-176-482-310	Sequence 310, Appl
27	88	6.4	596	9 US-10-176-757-310	Sequence 310, Appl
28	88	6.4	596	9 US-10-176-913-310	Sequence 310, Appl
29	88	6.4	596	9 US-10-180-552-310	Sequence 310, Appl
30	88	6.4	596	9 US-10-180-552-310	Sequence 310, Appl
31	88	6.4	596	9 US-10-173-700-310	Sequence 310, Appl
32	88	6.4	596	9 US-10-174-572-310	Sequence 310, Appl
33	88	6.4	596	9 US-10-174-579-310	Sequence 310, Appl
34	88	6.4	596	9 US-10-174-582-310	Sequence 310, Appl
35	88	6.4	596	9 US-10-174-588-310	Sequence 310, Appl
36	88	6.4	596	9 US-10-175-739-310	Sequence 310, Appl
37	88	6.4	596	9 US-10-175-740-310	Sequence 310, Appl
38	88	6.4	596	9 US-10-175-743-310	Sequence 310, Appl
39	88	6.4	596	9 US-10-176-488-310	Sequence 310, Appl
40	88	6.4	596	9 US-10-176-492-310	Sequence 310, Appl
41	88	6.4	596	9 US-10-176-747-310	Sequence 310, Appl
42	88	6.4	596	9 US-10-176-750-310	Sequence 310, Appl
43	88	6.4	596	9 US-10-176-985-310	Sequence 310, Appl
44	88	6.4	596	9 US-10-176-987-310	Sequence 310, Appl
45	88	6.4	596	9 US-10-176-991-310	Sequence 310, Appl

## ALIGNMENTS

## RESULT 1

US-09-862-027-27  
; Sequence 27, Application US/09862027  
; Patent No. US20020142428A1  
; GENERAL INFORMATION:  
; APPLICANT: Hodge, Martin R.  
; TITLE OF INVENTION: No. US20020142428A1el Kinases and Uses Thereof  
; FILE REFERENCE: 35800/234862  
; CURRENT APPLICATION NUMBER: US/09/862, 027  
; CURRENT FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: US 09/345,473  
; PRIOR FILING DATE: 1999-06-30  
; NUMBER OF SEQ ID NOS: 82  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 27  
; LENGTH: 669  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-862-027-27

Query Match 15.2%; Score 208.5; DB 10; Length 669;  
Best Local Similarity 27.7%; Pred. No. 3.1e-11;  
Matches 65; Conservative 46; Mismatches 87; Indels 37; Gaps 7;

QY 25 TSTNNVLTQTKDSAGNPVSSSAGRLVYSAPRLMED-SAVLTSPDTIINFETPTYSR 83  
Db 43 TNSGLMKLT-----NSSEFSYGHVFNPSVRFKNSPNTVSSFTTFVFAIVSNAL 96  
QY 84 IADGLAFIAPDSV-ISHGCGFLGLFPNANTLNNSSTSENOTTAKASSNVVAFEDTY 142  
Db 97 DGHGLAFVISPTKGLPYSSSSQYLGLF---NLTNNGD-----PSNHIVAFEDTF 143  
QY 143 LNPDYGGPNYTHIGDVNSIRSKVTAKWDQ-----NGKTATAHSNSVSK 189  
Db 144 QNQEFDMDNNHVGIDINLSSEKASTAGYEDDDGTGFKNIRLINOKPIQAWIEDSSRR 203  
QY 190 RLSVT---SYAGSKPATLSYDIELHTVLPBWVRVGLSASTGQDKERTNVHSWSF 241  
Db 204 QLNVTIHPHLPKPIPLLSLTLDLSPVLFDSMYVGFTSATGRLESSHYLWGTF 258

## RESULT 2

Qy	129	AASNVAVEPDTVLNPDYGDPNTH-----IGIDVNSIRSKYTAKWDWONGKIATAHI	182
Db	703	LPS---GLAFDSTNNITSGTPTNIGTSTIVSTDASGNKTTTFKYEVTRNSMSDSVS	758
Qy	183	SYNSVKRLSVTSYYAASKPATLSYDELHT----VLPEWRVGLSASTGDKERTVHS	238
Db	759	TSGTQQSQSVSTSKAQSQASTSTSSGIMTSTSAKTSKTSVSLSDSVSAKSILSTSES	818
Qy	239	WSFTSSLWTNVAKKENENKYITRGV	263
Db	819	NSVSSSTSTSLVNSQSVSMSGSV	843

```

Query Match          7.3%;  SCORE 99.5;  DB 10;  Length 639;
Best Local Similarity 23.7%;  Pred. No. 0.33;
Matches 57;  Conservative 43;  Mismatches 90;  Indels 51;  Gaps 13

QY 13 PNOEDLIFGHAATSNVNLQVTKLDSAGNPVSSACRVLVSAPLRWEDSAVLTSFDTII 72
      ||| :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
DB 403 PYPQD-----SDATNAHHITLGEPTSPNTVAFNSG--LHFA--RTFFHTSVLPDGSFTFI 454

```



```
FILE REFERENCE: 0342/IG548-US2
CURRENT APPLICATION NUMBER: US/09/893.519A
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: US 60/215,164
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: US 60/224,457
PRIOR FILING DATE: 2000-08-10
NUMBER OF SEQ ID NOS: 146
SOFTWARE: PatentIn version 3.1
SEQ ID NO 41
LENGTH: 415
TYPE: PRT
ORGANISM: Candida albicans
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Corresponds to SEQ ID NO: 114
US-09-893-519A-41

Query Match 6.7%; Score 92; DB 9; Length 415;
Best Local Similarity 23.7%; Pred. No. 0.92;
Matches 54; Conservative 30; Mismatches 100; Indels 44; Gaps 9;

Qy 3 SLSFSTKFPNQEDLIFQGHATSTNNVQVTKLDSAGNPVSSSAGRVLYSAPRLWEDS 62
Db 17 SISGQLKFDSDIAPYIKELTEKEN---VRKVPFSGNTIGIEASKAL-----SE 64

Qy 63 AVLTSEFDIINEFISTPTYSRIADGLAFPIAPDSVISYHGGFLGLFPNANTLNNSSTSE 122
Db 65 ALLKHKDTIVEINFSDLYTGRLNETI-----POS-LEYLLPALSKLPNUKLINLSDNAF 117

Qy 123 NOTT---TKAASNNVAVFDTYLPNDYDGPVNIHIDV-----NSIRSKVT 167
Db 118 GLQTDPIEAYLAKAVSIEHLILNNGMPFAGSRIGGSLFKLAKAKAEGKESLTKFIC 177

Qy 168 AKWDQWNGKIATAHISYNSVS-----KRLSVTSYYA-GSKPATLSYDIE 210
Db 178 GRNRLNGS-----VNYLSVGLRNHDKLEVRVLYQNGIRPAGISKILVE 220

RESULT 8
US-09-841-132-394
; Sequence 394, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 394
; LENGTH: 1723
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-841-132-394

Query Match 6.7%; Score 91.5; DB 10; Length 1723;
Best Local Similarity 23.5%; Pred. No. 6.9;
Matches 63; Conservative 32; Mismatches 106; Indels 67; Gaps 12;

Qy 14 NORDLIFOGHATSTNNVQVTKLDSAGNPVSSSA---GRVLYSAPRLWEDSAVLTSFDT 70
Db 1007 NPEITFKGNYATGGAIGCIDITNGSPPRKVSADNGSVLF-----QDNSALNRGGA 1059

Qy 71 IINFEISTPTYSRIADGLAFPIAPDSVISYHGG-----FLGLFPNANTL-NNSTSE 122
Db 1060 IYGETIDISRTG-----ATFIGNS-----SKHDGSAICCSALTALPANSQILFENKNKYTE 1109

Qy 123 NOTTTKAASNNVAVFDTYLPNDYDGPVNIHIDVNSIRSKVTAKWDQWNGKI----- 177

FILE REFERENCE: 0342/IG548-US2
CURRENT APPLICATION NUMBER: US/09/893.519A
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: US 60/215,164
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: US 60/224,457
PRIOR FILING DATE: 2000-08-10
NUMBER OF SEQ ID NOS: 146
SOFTWARE: PatentIn version 3.1
SEQ ID NO 41
LENGTH: 415
TYPE: PRT
ORGANISM: Candida albicans
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Corresponds to SEQ ID NO: 114
US-09-893-519A-41

Query Match 6.7%; Score 91; DB 10; Length 433;
Best Local Similarity 23.0%; Pred. No. 1.2;
Matches 70; Conservative 37; Mismatches 108; Indels 90; Gaps 15;

Qy 24 ATSTNNVQVTKLDSAG-----NPVSSSAGRVLYSAPRLWEDSAVLTSFD 69
Db 8 ATSKNOVLQRRPLESTNSNHGFASSLQAIPEMTMSGSDNASFQSLPLSM--SSSQSTSS 65

Qy 70 TIINFEISTP-YTSRIADGL-AFFIAPDPSVISYHGGFLG-----LFPNAN 113
Db 66 RRENPNAPPEYTDRADEIKRLLASSPSRRSHSSSMHSASRRSSVAESGSLSDNAS 125

Qy 114 TLNNS-----STSENQTTTKAASN-----VVAVEFDYTL-----NPD--YGDPNY 152
Db 126 SYQSSIFSAPSTVHTQLTNDSSFEFPNHLITRVSLEALPKTFYDMYSPDILLADPSN 185

Qy 153 I-----HIGIDVNSIRS-----KVAKWDQWNGKIATAHISYNSVSKRLSVTS 195
Db 186 ILCNGRPFKTKRELDWDLNDLRSLIVLEKLPWGNQQLPEVIT--VGDMPQFRQLQLP 243

Qy 196 YVAGSKPATLSYDIEIHLTVLPWVRVGLSASTQDQKERNVTSHWSFTSSLTWTVAKKENE 255
Db 244 LYSDE-----TIATLVHSDLY-----MEANLDYEFKLTSAKYT-VATARKR 285

Qy 256 NKYIT 260
Db 286 HEHIT 290
```



RESULT 10  
US-10-186-399-3  
; Sequence 3, Application US/10186399  
; Patent No. US20020173481A1  
; GENERAL INFORMATION:  
; APPLICANT: Ekman, Niklas  
; APPLICANT: Arighi, Elena  
; APPLICANT: Vastrik, Imre  
; APPLICANT: Tamagnone, Luca  
; APPLICANT: Alitalo, Kari  
; TITLE OF INVENTION: REGULATION OF VASCULAR ENDOTHELIUM USING BMX TYROSINE  
; FILE REFERENCE: 28113/31941A  
; CURRENT FILING DATE: 2002-07-01  
; PRIOR APPLICATION NUMBER: US/10/186,399  
; PRIOR FILING DATE: 2002-07-01  
; PRIOR APPLICATION NUMBER: US 08/320,432  
; PRIOR FILING DATE: 1994-10-07  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 675  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-186-399-3

Query Match  
Best Local Similarity 23.0%; Score 91; DB 9; Length 675;  
Matches 56; Conservative 44; Mismatches 72; Indels 72; Gaps 16;  
QY 72 INFESTP-----YTSRIA--DGLAFFIAP-----PDSVISYHGGFL 106  
: : | | | | : | : | | | : | : | | | : | : | | | : | : | | | :  
Db 66 VNLDEQTPVERQYPFQIVYKDGLLYVYASNEESRSQWLKALQKEIRGNPHLLVYHSGF- 124  
QY 107 GLFPNANTL---NNSSTSENQTTTKAASSNVAVFDTYLNPDYGD-----PNVIHIGID 158  
: : | | | | : | : | | | : | : | | | : | : | | | : | : | | | :  
Db 125 --FVDGKFLCCQSCKAAPGCTLWEAYANLHTAVNEEKHRVPTFDRVLKIPRAVPV-LK 181  
QY 159 VNSIRSKVT-AKWDWONGK-----IATAHISYNSVSKRLSVTSYYAGSKPATLSYDI 209  
: : | | | | : | : | | | : | : | | | : | : | | | : | : | | | :  
Db 182 MDAPSSSTTLAQYDNESKKNYGSQPPSSSTSLAQYDSNSKKI-----YGSQP---NFNM 232  
QY 210 EL--HTVLPW--VRVGLSASTGQD-----KERNIVH-----SWFTSSLWTNVAKEN 254  
: : | | | | : | : | | | : | : | | | : | : | | | : | : | | | :  
Db 233 QYIPREDPDMQVVRKLKSSSESVASSNQKERNVNHHTTSKISWEPPES--SSSEEEEN 290  
QY 255 ENKY 258  
: : |  
Db 291 LDDY 294

RESULT 11  
US-09-977-260-4  
; Sequence 4, Application US/09977260  
; Publication No. US20020192790A1  
; GENERAL INFORMATION:  
; APPLICANT: ULLRICH, AXEL  
; APPLICANT: GISHIZKY, MIKHAIL  
; APPLICANT: SURES, IRMINGARD  
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
; FILE REFERENCE: 038602/1260  
; CURRENT APPLICATION NUMBER: US/09/977,260  
; CURRENT FILING DATE: 2001-10-16  
; PRIOR APPLICATION NUMBER: 08/232,545  
; PRIOR FILING DATE: 1994-04-22  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 675  
; TYPE: PRT  
; ORGANISM: Unknown Organism  
; FEATURE:  
US-09-977-260-4

; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte  
; OTHER INFORMATION: kinase 2  
US-09-977-260-4

Query Match  
Best Local Similarity 23.0%; Score 91; DB 9; Length 675;  
Matches 56; Conservative 44; Mismatches 72; Indels 72; Gaps 16;  
QY 72 INFESTP-----YTSRIA--DGLAFFIAP-----PDSVISYHGGFL 106  
: : | | | | : | : | | | : | : | | | : | : | | | : | : | | | :  
Db 66 VNLDEQTPVERQYPFQIVYKDGLLYVYASNEESRSQWLKALQKEIRGNPHLLVYHSGF- 124  
QY 107 GLFPNANTL---NNSSTSENQTTTKAASSNVAVFDTYLNPDYGD-----PNVIHIGID 158  
: : | | | | : | : | | | : | : | | | : | : | | | : | : | | | :  
Db 125 --FVDGKFLCCQSCKAAPGCTLWEAYANLHTAVNEEKHRVPTFDRVLKIPRAVPV-LK 181  
QY 159 VNSIRSKVT-AKWDWONGK-----IATAHISYNSVSKRLSVTSYYAGSKPATLSYDI 209  
: : | | | | : | : | | | : | : | | | : | : | | | : | : | | | :  
Db 182 MDAPSSSTTLAQYDNESKKNYGSQPPSSSTSLAQYDSNSKKI-----YGSQP---NFNM 232  
QY 210 EL--HTVLPW--VRVGLSASTGQD-----KERNIVH-----SWFTSSLWTNVAKEN 254  
: : | | | | : | : | | | : | : | | | : | : | | | : | : | | | :  
Db 233 QYIPREDPDMQVVRKLKSSSESVASSNQKERNVNHHTTSKISWEPPES--SSSEEEEN 290  
QY 255 ENKY 258  
: : |  
Db 291 LDDY 294

RESULT 12  
US-09-977-269-4  
; Sequence 4, Application US/09977269  
; Patent No. US20020082037A1  
; GENERAL INFORMATION:  
; APPLICANT: ULLRICH, AXEL  
; APPLICANT: GISHIZKY, MIKHAIL  
; APPLICANT: SURES, IRMINGARD  
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
; FILE REFERENCE: 038602/1260  
; CURRENT APPLICATION NUMBER: US/09/977,269  
; CURRENT FILING DATE: 2001-10-16  
; PRIOR APPLICATION NUMBER: 08/232,545  
; PRIOR FILING DATE: 1994-04-22  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 675  
; TYPE: PRT  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte  
; OTHER INFORMATION: kinase 2  
US-09-977-269-4

Query Match  
Best Local Similarity 23.0%; Score 91; DB 10; Length 675;  
Matches 56; Conservative 44; Mismatches 72; Indels 72; Gaps 16;  
QY 72 INFESTP-----YTSRIA--DGLAFFIAP-----PDSVISYHGGFL 106  
: : | | | | : | : | | | : | : | | | : | : | | | : | : | | | :  
Db 66 VNLDEQTPVERQYPFQIVYKDGLLYVYASNEESRSQWLKALQKEIRGNPHLLVYHSGF- 124  
QY 107 GLFPNANTL---NNSSTSENQTTTKAASSNVAVFDTYLNPDYGD-----PNVIHIGID 158  
: : | | | | : | : | | | : | : | | | : | : | | | : | : | | | :  
Db 125 --FVDGKFLCCQSCKAAPGCTLWEAYANLHTAVNEEKHRVPTFDRVLKIPRAVPV-LK 181  
QY 159 VNSIRSKVT-AKWDWONGK-----IATAHISYNSVSKRLSVTSYYAGSKPATLSYDI 209  
: : | | | | : | : | | | : | : | | | : | : | | | : | : | | | :  
Db 182 MDAPSSSTTLAQYDNESKKNYGSQPPSSSTSLAQYDSNSKKI-----YGSQP---NFNM 232  
QY 210 EL--HTVLPW--VRVGLSASTGQD-----KERNIVH-----SWFTSSLWTNVAKEN 254  
: : | | | | : | : | | | : | : | | | : | : | | | : | : | | | :  
Db 233 QYIPREDPDMQVVRKLKSSSESVASSNQKERNVNHHTTSKISWEPPES--SSSEEEEN 290

QY 255 ENKY 258  
:  
Db 291 LDDY 294

RESULT 13  
US-09-841-132-395  
; Sequence 395, Application US/09841132  
; Patent No. US2002006184A1  
; GENERAL INFORMATION:  
; APPLICANT: Bhatia, Ajay  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Probst, Peter  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; FILE REFERENCE: 210121.469C8  
; CURRENT APPLICATION NUMBER: US/09/841,132  
; CURRENT FILING DATE: 2001-04-23  
; NUMBER OF SEQ ID NOS: 599  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 395  
; LENGTH: 1723  
; TYPE: PRT  
; ORGANISM: Chlamydia pneumoniae  
US-09-841-132-395

Query Match 6.6%; Score 90.5; DB 10; Length 1723;  
Best Local Similarity 23.1%; Pred. No. 8;  
Matches 62; Conservative 33; Mismatches 106; Indels 67; Gaps 12;

QY 14 NQEDLIFOGHATSTNNVLOVTKLDASAGNPVSSA---GRVLVSAPRLRWEDSAVLTSPTD 70  
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
Db 1007 NNPEITFFGNVAETGAIGCIDLTNGSPPKVSIADNGSVLF-----QDSALNRGGA 1059  
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
QY 71 IINFEISTPYTSRIADGLAFFIAPPDSVISYHG-----FLGLFPNANTL--NNSSTSE 122  
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
Db 1060 IYGETIDISRTG-----ATFIGNS----SKHDGSAICCSALTALPNSQLFENKKYTE 1109  
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
QY 123 NQTTHKAASSNVAVFDTYLPNDYGDPNYIHIGIDVNSRSTKVTAKWQNGKI----- 177  
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
Db 1110 TTATTTKASINILGAA-----IYG-----NNETSITISLSAENGSIFFKN 1150  
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
QY 178 -ATAHISVNSVKRSLSVTSYAGSKPATLSYDIELHTVLPWVRVGLSASTGODKERNTV 236  
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
Db 1151 LCTATNKVCSTAGNVKFTAEIASAGKAISFYDA-----VVWSIKETNAQELKLN- 1200  
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
QY 237 HSWFTSSLWTNVAKKENPNK-YITRGV 263  
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|  
Db 1201 ---KATGTGTTLFSGELHENKSYPQKV 1225  
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 14  
US-09-886-468-20  
; Sequence 20, Application US/09886468  
; Patent No. US20020037293A1  
; GENERAL INFORMATION:  
; APPLICANT: Aventis Pasteur Limited  
; TITLE OF INVENTION: Chlamydia antigens and corresponding DNA fragments and uses there  
; FILE REFERENCE: 77813-5  
; CURRENT APPLICATION NUMBER: US/09/886,468  
; CURRENT FILING DATE: 1999-12-23  
; PRIOR APPLICATION NUMBER: 60/113,280  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/113,281  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/113,282  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/113,283  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/113,284  
; PRIOR FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 60/113,285

Query Match 6.5%; Score 89.5; DB 10; Length 632;  
Best Local Similarity 21.1%; Pred. No. 2.8;  
Matches 58; Conservative 37; Mismatches 97; Indels 83; Gaps 15;

QY 11 FDPNQEDLIFOCHAT-STNNVLQV-----TKLDSAGNPVSSSAGRVLYSAPRLWE 60  
Db 210 FDLNOKYTFIIGSATGAANNKHQIGVTLFEAYFTKPTTEANPDVIELG----- 257

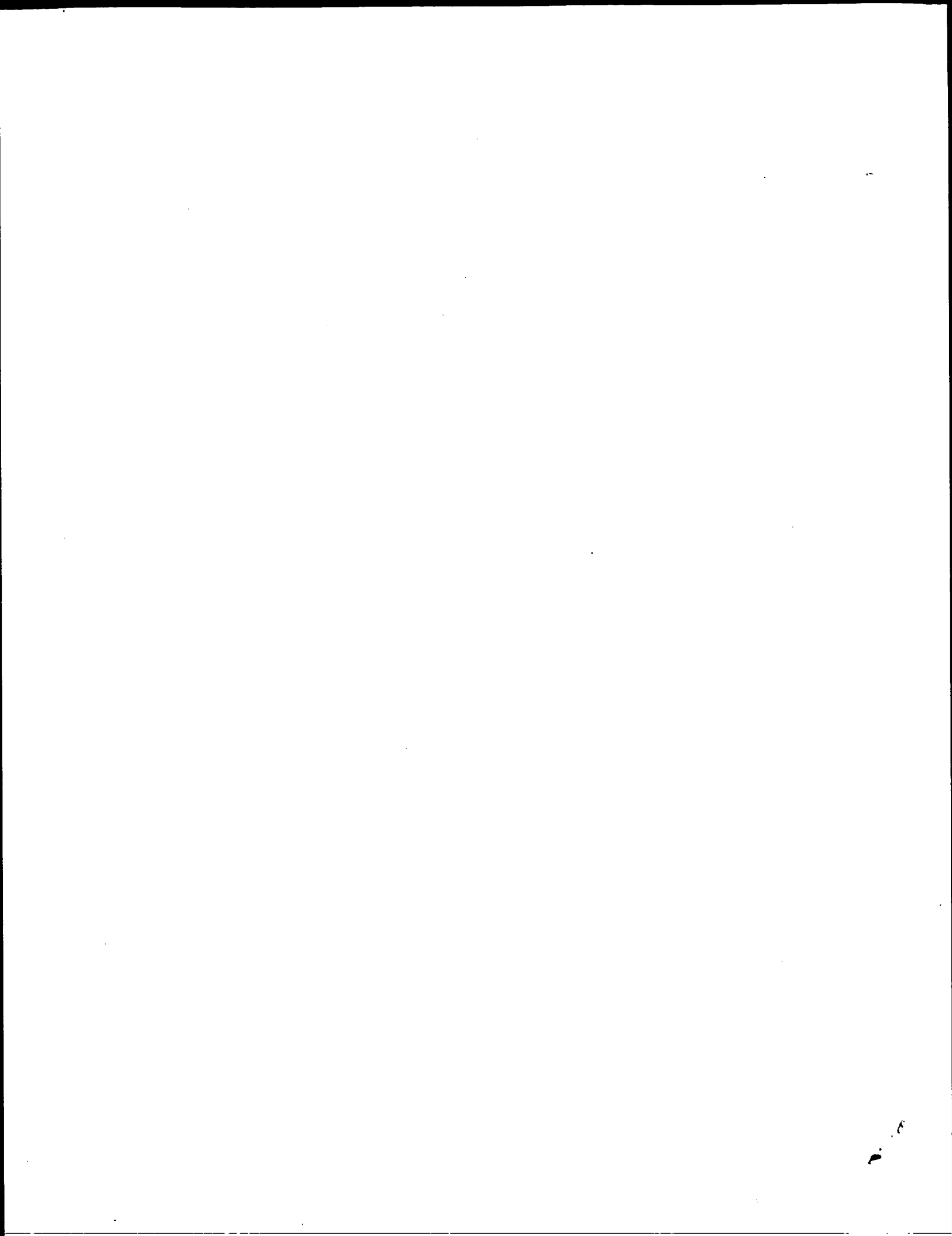
QY 61 DSAVLTSFDTLIINFEISTPYTSRIADGLAFFIAPDPSVISYHGGFLGLFPNANTLNNSST 120  
Db 258 -----TAFDPLNHPEPIGLKATDEVDGDI-----KDITVEF-----NDIDTSKP 296

QY 121 SENOTTTKAAS-----SNVAVVEFDYTLNPDYG-----DP-NYIHIGIDV----- 159  
Db 297 GAYRVYKVVNSYGESDEKTEIEVVYTKPTITAHDTIKKDLAFDPLNVEPIGLKATDPI 356

QY 160 -NSIRSKVTAKW-DWQNGKIATAHLSYNSVS-----KRLSVTSYYAGSKPATLSYDIE 210  
Db 357 DGDITDKIAVKFNVDTSKPGKYHVYKVINSYEKIDEKTEIVTVY---TKPSIVAH DVE 413

QY 211 L--HTVLP--EWVRVGLSAS---TGQDKERNVTVHS 238  
Db 414 IKKDTAFDPLNVEPIGLKATDPIDGDTDKITVES 448

Search completed: February 26, 2003, 16:52:28  
Job time : 26.2285 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:40:30 ; Search time 32.2996 Seconds  
(without alignments)  
1089.120 Million cell updates/sec

Title: US-09-476-485A-2

Perfect score: 1370

Sequence: 1 AGSLGFSFTKFDPNQEDLIF.....LMTNVAKENENKVIITRGVL 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1362	99.4	264	20	AAW87973
2	1362	99.4	264	22	AAG62890
3	1362	99.4	286	22	AAG62894
4	800.5	58.4	234	22	AAG62901
5	797	58.2	303	22	AAG62898
6	642	46.9	290	21	AAV58736
7	637	46.5	290	21	AAV58737
8	552	40.3	237	16	AAV4765
9	463	33.8	285	20	AAV06812
10	441	32.2	285	20	AAV06814

11	440	32.1	285	20	AAV06811
12	439	32.0	285	20	AAV06815
13	422	30.8	275	21	AAV58738
14	420	30.7	242	15	AAV45911
15	420	30.7	262	20	AAV27265
16	413	30.1	242	15	AAV45912
17	378.5	27.6	239	16	AAV47667
18	366	26.7	228	16	AAV47666
19	345	25.2	236	16	AAV4807
20	321.5	23.5	274	20	AAV06813
21	295	21.5	681	23	ABV93833
22	288	21.0	715	23	ABV92849
23	270.5	19.7	282	21	AAV25452
24	264	19.3	265	10	AAV91967
25	259	18.9	711	23	ABV93337
26	256.5	18.7	691	23	ABV93681
27	244	17.8	632	21	AAV25109
28	241	17.6	652	23	ABV93374
29	229.5	16.8	677	23	ABV92654
30	223.5	16.3	283	23	ABV91344
31	218.5	15.9	272	23	ABV91343
32	217	15.8	674	23	ABV93878
33	214.5	15.7	675	23	ABV92029
34	213	15.5	244	10	AAV93640
35	208.5	15.2	669	23	ABV93142
36	201.5	14.7	688	23	ABV93318
37	194.5	14.2	693	23	ABV92247
38	188.5	13.8	616	23	ABV93900
39	188.5	13.8	649	23	ABV9134
40	186	13.6	242	23	ABV91342
41	185	13.5	658	23	ABV92107
42	184.5	13.5	685	23	ABV93319
43	181	13.2	627	23	ABV91918
44	176	12.8	667	23	ABV92524
45	174.5	12.7	623	23	ABV91919

## ALIGNMENTS

### RESULT 1

AAW87973

ID AAW87973 standard; Protein; 264 AA.

XX AAW87973;

AC AAW87973;

XX 13-APR-1999 (first entry)

DT A lectin derived progenitor cell preservation factor.

DE Lectin derived progenitor cell preservation factor; progenitor cell;

KW haematopoietic cell; cultured cell preservation; anticancer therapy;

KW myeloblastic therapy; sickle-cell anaemia; ablative therapy protection;

KW FLK2/FLT3 receptor.

XX Dolichos lab lab.

OS WO9859038-A1.

XX 30-DEC-1998.

PD 23-JUN-1998; 98WO-US13046.

XX 24-JUN-1997; 97US-0881189.

XX (IMCL-) IMCLONE SYSTEMS INC.

PA (REGC) UNIV CALIFORNIA.

XX Chrispeels MJ, Colucci MG, Moore JG;

XX WPI: 1999-081274/07.

DR N-PSDB; AAX03593.

XX

PT New nucleic acid encoding plant lectin that preserves progenitor  
PT cells - particularly haematopoietic progenitors, useful for bone  
PT marrow reconstitution after ablative therapy, and to increase DNA  
PT transfer in gene therapy  
XX  
PS Claim 1; Page 30-31; 72pp; English.  
XX  
CC The present sequence represents a lectin derived progenitor cell  
CC preservation factor. The protein is used to preserve unipotent,  
CC pluripotent or totipotent progenitor cells, especially haematopoietic  
CC cells, and also progenitors from nerve, muscle, skin, gut, bone,  
CC kidney, liver, pancreas or thymus. Specific applications are  
CC preservation of cultured cells intended for administration after  
CC (anticancer) myeloablative therapy (bone marrow or whole-body irradiation  
CC or chemotherapy) to reconstitute the haematopoietic system; enrichment  
CC of progenitor cells (e.g. during ex vivo purging of malignant cells);  
CC treatment of tissues containing haematopoietic progenitors for subsequent  
CC transplant to improve haematopoietic competence; improving transfer of  
CC exogenous DNA to progenitor cells (in gene therapy of various  
CC haematological disorders, e.g. sickle-cell anaemia); and protection  
CC against ablative therapy (to eliminate proliferating cells specifically),  
CC followed by re-establishment of differentiation and proliferation of  
CC preserved progenitors. The protein, when linked to magnetic beads, may  
CC also be used to isolate cells that express the FLK2/FLT3 receptor.  
XX  
XX Sequence 264 AA;  
SQ  
Query Match 99.4%; Score 1362; DB 20; Length 264;  
Best Local Similarity 99.6%; Pred. No. 3.le-124;  
Matches 263; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 AGSLSFSTKFPDNOEDLIFQGHATSTNNVLOVTKLDSAGNPVSSSAGRVLYSAPRLWE 60  
DB 1 AQSLSFSTKFPDNOEDLIFQGHATSTNNVLOVTKLDSAGNPVSSSAGRVLYSAPRLWE 60  
QY 61 DSAVLTSFDTIINFELSTPYTSRIADGLAFFIAPPDPSVISYHGGLGLFPNANTLNSST 120  
DB 61 DSAVLTSFDTIINFELSTPYTSRIADGLAFFIAPPDPSVISYHGGLGLFPNANTLNSST 120  
QY 121 SENQTTTAAASNNAVVEFDTYLNPDPYIHIGIDVNSIRSKVTAKWDQWNGKIATA 180  
DB 121 SENQTTTAAASNNAVVEFDTYLNPDPYIHIGIDVNSIRSKVTAKWDQWNGKIATA 180  
QY 181 HTSYNSVKRLSVTSYYAGSKPATLSYDIELHTVLPWVRVGLSASTGQDKERTVHWS 240  
DB 181 HTSYNSVKRLSVTSYYAGSKPATLSYDIELHTVLPWVRVGLSASTGQDKERTVHWS 240  
QY 241 FTSSLTWNVAKKENENKYITRGVL 264  
DB 241 FTSSLTWNVAKKENENKYITRGVL 264  
RESULT 2  
AAG62890  
ID AAG62890 standard; Protein; 264 AA.  
XX  
XX AAG62890;  
XX  
XX 17-SEP-2001 (first entry)  
XX  
XX Amino acid sequence of a hyacinth bean FRIL polypeptide.  
XX  
XX FRIL; FLK2/FLT3 tyrosine kinase receptor-interacting lectin;  
XX progenitor cell preservation factor; radiotherapy; chemotherapy;  
KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
KW severe combined immunodeficiency; aplastic anemia; tissue repair.  
XX  
XX Dolichos lab lab.  
OS  
XX  
XX WO200149851-A1.  
PN  
XX  
XX 12-JUL-2001.  
PD  
XX

BF 30-DEC-1999; 99WO-US31307.  
XX  
PR 30-DEC-1999; 99WO-US31307.  
XX  
PA (PHYL-) PHYLOGIX LLC.  
XX  
PI Colucci MG, Chrispeels MJ, Moore JG;  
XX  
XX WPI; 2001-441882/47.  
DR N-PSDB; AAH42287.  
DR  
XX  
XX Legume progenitor cell preservation factors for in vivo or ex vivo  
PT preservation of hematopoietic progenitor cells and as therapeutics for  
PT alleviating/reducing progenitor cell-depleting activity of cancer  
PT therapeutics  
XX  
PS Example 1; Page 54-55; 173pp; English.  
XX  
CC The present sequence represents a FRIL (FLK2/FLT3 tyrosine kinase  
CC receptor-interacting lectin) polypeptide. The specification describes a  
CC composition of one or more members of FRIL family of progenitor cell  
CC preservation factors. The composition is useful for alleviating or  
CC reducing the hematopoietic progenitor cell-depleting activity of  
CC a therapeutic treatment, including radiotherapeutic and/or  
CC chemotherapeutic treatments. Administration of FRIL compositions to a  
CC patient prior to treatment of the patient with a therapeutic treatment  
CC having a hematopoietic progenitor cell-depleting activity alleviates or  
CC reduces the hematopoietic progenitor cell-depleting activity of the  
CC therapeutic treatment in the patient. FRIL family members are useful for  
CC isolating population of progenitor cells, hemangioblasts, and mesenchymal  
CC stem cells. The composition is administered to reduce progenitor cell  
CC depleting effects of chemotherapeutics, so that the patient can receive  
CC a higher dose of the chemotherapeutic and preferably recover from cancer.  
CC It is also administered to patients having, or predisposed to developing  
CC a condition where the patients hematopoietic progenitor cells are  
CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
CC The isolated mesenchymal cells are useful for tissue repair.  
XX  
XX Sequence 264 AA;  
SQ  
Query Match 99.4%; Score 1362; DB 22; Length 264;  
Best Local Similarity 99.6%; Pred. No. 3.le-124;  
Matches 263; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 AGSLSFSTKFPDNOEDLIFQGHATSTNNVLOVTKLDSAGNPVSSSAGRVLYSAPRLWE 60  
DB 1 AQSLSFSTKFPDNOEDLIFQGHATSTNNVLOVTKLDSAGNPVSSSAGRVLYSAPRLWE 60  
QY 61 DSAVLTSFDTIINFELSTPYTSRIADGLAFFIAPPDPSVISYHGGLGLFPNANTLNSST 120  
DB 61 DSAVLTSFDTIINFELSTPYTSRIADGLAFFIAPPDPSVISYHGGLGLFPNANTLNSST 120  
QY 121 SENQTTTAAASNNAVVEFDTYLNPDPYIHIGIDVNSIRSKVTAKWDQWNGKIATA 180  
DB 121 SENQTTTAAASNNAVVEFDTYLNPDPYIHIGIDVNSIRSKVTAKWDQWNGKIATA 180  
QY 181 HTSYNSVKRLSVTSYYAGSKPATLSYDIELHTVLPWVRVGLSASTGQDKERTVHWS 240  
DB 181 HTSYNSVKRLSVTSYYAGSKPATLSYDIELHTVLPWVRVGLSASTGQDKERTVHWS 240  
QY 241 FTSSLTWNVAKKENENKYITRGVL 264  
DB 241 FTSSLTWNVAKKENENKYITRGVL 264  
RESULT 3  
AAG62894  
ID AAG62894 standard; Protein; 286 AA.  
XX  
XX AAG62894;  
XX  
XX 17-SEP-2001 (first entry)  
DT  
XX

DE Alpha-amylase inhibitor signal peptide and FRIL fusion.  
 KW FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;  
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;  
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
 KW severe combined immunodeficiency; aplastic anemia; tissue repair;  
 KW alpha-amylase inhibitor gene.  
 XX  
 OS Synthetic.  
 OS Unidentified.  
 OS Dolichos lab lab.  
 XX  
 PN WO200149851-A1.  
 XX  
 XX 12-JUL-2001.  
 XX  
 XX 30-DEC-1999; 99WO-US31307.  
 XX  
 XX 30-DEC-1999; 99WO-US31307.  
 XX  
 PA (PHYL-) PHYLOGIX LLC.  
 XX  
 PI Colucci MG, Chrispeels MJ, Moore JG;  
 XX  
 DR WPI; 2001-441882/47.  
 DR N-PSDB; AAH42295.  
 XX  
 PT Legume Progenitor cell preservation factors for in vivo or ex vivo  
 PT preservation of hematopoietic progenitor cells and as therapeutics for  
 PT alleviating/reducing progenitor cell-depleting activity of cancer  
 PT therapeutics  
 XX  
 PS Example 1; Page 59; 173pp; English.  
 XX  
 CC The present sequence represents fusion protein of alpha-amylase inhibitor  
 CC signal peptide and FRIL (Flk2/Flt3 tyrosine kinase receptor-interacting  
 CC lectin). The specification describes a composition of one or more members  
 CC of FRIL family of progenitor cell preservation factors. The composition  
 CC is useful for alleviating or reducing the hematopoietic progenitor  
 CC cell-depleting activity of a therapeutic treatment, including  
 CC radiotherapeutic and/or chemotherapeutic treatments. Administration of  
 CC FRIL compositions to a patient prior to treatment of the patient with  
 CC a therapeutic treatment having a hematopoietic progenitor cell-depleting  
 CC activity alleviates or reduces the hematopoietic progenitor  
 CC cell-depleting activity of the therapeutic treatment in the patient.  
 CC FRIL family members are useful for isolating population of progenitor  
 CC cells, hemangioblasts, and mesenchymal stem cells. The composition is  
 CC administered to reduce progenitor cell depleting effects of  
 CC chemotherapeutics, so that the patient can receive a higher dose of the  
 CC chemotherapeutic and preferably recover from cancer. It is also  
 CC administered to patients having, or predisposed to developing a  
 CC condition where the patients' hematopoietic progenitor cells are  
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
 CC The isolated mesenchymal cells are useful for tissue repair.  
 XX  
 SQ Sequence 286 AA;  
 Query Match 99.4%; Score 1362; DB 22; Length 286;  
 Best Local Similarity 99.6%; Pred. No. 3.5e-124;  
 Matches 263; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 AGSLFSFTKPDNQBDLFIQGHATSTNNVQVTKLDSAGNPVSSSAGRVLYSAPLRLWE 60  
 DB 23 AQSLFSFTKPDNQBDLFIQGHATSTNNVQVTKLDSAGNPVSSSAGRVLYSAPLRLWE 82  
 OY 61 DSAVLTSFTDIINFEISTPYTRIADGLAFFIAPDPSVSIYHGGFLGLFPNANTLNSST 120  
 DB 83 DSAVLTSFTDIINFEISTPYTRIADGLAFFIAPDPSVSIYHGGFLGLFPNANTLNSST 142  
 OY 121 SENQTTTAAASNVAVEFDYLNPDYGPDPNYTHIGIDVNSIRSKVTAKWDWONGKIATA 180  
 DB 143 SENQTTTAAASNVAVEFDYLNPDYGPDPNYTHIGIDVNSIRSKVTAKWDWONGKIATA 202

OY 181 HISYNSVSKRLSVTSYYAGSKPATLSYDIELHTVLPWVRVGLSASTGDKERNTVHSWS 240  
 DB 203 HISYNSVSKRLSVTSYYAGSKPATLSYDIELHTVLPWVRVGLSASTGDKERNTVHSWS 262  
 OY 241 FTSSLWNTNVAKENKENKYITRGVL 264  
 DB 263 FTSSLWNTNVAKENKENKYITRGVL 286  
 RESULT 4  
 AAG62901  
 ID AAG62901 standard; Protein; 234 AA.  
 XX  
 AC AAG62901;  
 XX  
 DT 17-SEP-2001 (first entry)  
 XX  
 DE Amino acid sequence of a partial FRIL polypeptide.  
 XX  
 KW FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;  
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;  
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
 KW severe combined immunodeficiency; aplastic anemia; tissue repair.  
 XX  
 OS Sphenostylis stenocarpa.  
 XX  
 PN WO200149851-A1.  
 XX  
 PD 12-JUL-2001.  
 XX  
 PF 30-DEC-1999; 99WO-US31307.  
 XX  
 PR 30-DEC-1999; 99WO-US31307.  
 XX  
 PA (PHYL-) PHYLOGIX LLC.  
 XX  
 PI Colucci MG, Chrispeels MJ, Moore JG;  
 XX  
 DR WPI; 2001-441882/47.  
 XX

Legume Progenitor cell preservation factors for in vivo or ex vivo  
 preservation of hematopoietic progenitor cells and as therapeutics for  
 alleviating/reducing progenitor cell-depleting activity of cancer  
 therapeutics  
 Example 22; Page 116-117; 173pp; English.

The present sequence is a partial a FRIL (Flk2/Flt3 tyrosine kinase  
 receptor-interacting lectin) protein. The specification describes a  
 composition of one or more members of FRIL family of progenitor cell  
 preservation factors. The composition is useful for alleviating or  
 reducing the hematopoietic progenitor cell-depleting activity of a  
 therapeutic treatment, including radiotherapeutic and/or  
 chemotherapeutic treatments. Administration of FRIL compositions to a  
 patient prior to treatment of the patient with a therapeutic treatment  
 having a hematopoietic progenitor cell-depleting activity alleviates or  
 reduces the hematopoietic progenitor cell-depleting activity of the  
 therapeutic treatment in the patient. FRIL family members are useful for  
 isolating population of progenitor cells, hemangioblasts, and mesenchymal  
 stem cells. The composition is administered to reduce progenitor cell  
 depleting effects of chemotherapeutics, so that the patient can receive  
 a higher dose of the chemotherapeutic and preferably recover from cancer.  
 It is also administered to patients having, or predisposed to developing  
 a condition where the patients' hematopoietic progenitor cells are  
 depleted, such as severe combined immunodeficiency or aplastic anemia.  
 The isolated mesenchymal cells are useful for tissue repair.

SQ Sequence 234 AA;

Query Match 58.4%; Score 800.5; DB 22; Length 234;  
 Best Local Similarity 66.7%; Pred. No. 1.3e-69;  
 Matches 156; Conservative 30; Mismatches 45; Indels 3; Gaps 2;

QY 1 AGSLSFSTKFDPNQEDLIFOGHATSTNNVLOVTKLDSAGNPVSSAGRVLYSAPLRLWE 60  
DQ 1 AQSLSFSTKFDSDQKDLMFQGHATSTSSNIOVTKLDSAGNPVSSAGRVLYSAPLRLWE 60  
QY 61 DSAVLTSFDTTINFEISPTYSRIADGLAFFIAPPDSVI--SYHGGFLGLFPNANTLNS 118  
DQ 61 SSVVSTFETTTFOISPTYSPPGDLGAFIAPVDTVPDPSAGNLLGLFPNANTLNS 120  
QY 119 STSENQTTTAAASNNVAVFEDTYLNDYDGNPIHIGIDVNSIRSKVTAKWDQNGKI 177  
DQ 121 TTSKETTIDVNAASNNVAVFEDTYLNDYDGNPIHIGIDVNSIRSKVTAKWDQNGKI 180  
QY 178 ATAHISYNSVSKRLSVTSYYAGSKPATLSYDIELHTVLPWVVRVGLSASTGQDK 231  
DQ 181 ATAHISYNSASRLSVTTPYGGKAVSLSHDVELTQVLPQWIRVGSASTGLEK 234  
RESULT 5  
AAG62898 standard; Protein; 303 AA.  
XX AC AAG62898;  
XX DT 17-SEP-2001 (first entry)  
XX DE Amino acid sequence of a french bean FRIL polypeptide.  
XX KW FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;  
KW progenitor cell preservation factor; radiotherapy; chemotherapy;  
KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
KW severe combined immunodeficiency; aplastic anemia; tissue repair.  
XX OS Phaseolus vulgaris.  
XX PN WO200149851-A1.  
XX PD 12-JUL-2001.  
XX PF 30-DEC-1999; 99WO-US31307.  
XX PR 30-DEC-1999; 99WO-US31307.  
XX PA (PHYL-) PHYLOGIX LLC.  
XX PI Colucci MG, Chrispeels MJ, Moore JG;  
XX DR WPI: 2001-441882/47.  
XX DR N-PSDB; AAH42306.  
XX Legume Progenitor cell preservation factors for in vivo or ex vivo  
PT preservation of hematopoietic progenitor cells and as therapeutics for  
PT alleviating/reducing progenitor cell-depleting activity of cancer  
PT therapeutics  
XX Example 5; Page 81; 173pp; English.  
XX The present sequence represents a FRIL (Flk2/Flt3 tyrosine kinase  
CC receptor-interacting lectin) polypeptide. The specification describes a  
CC composition of one or more members of FRIL family of progenitor cell  
CC preservation factors. The composition is useful for alleviating or  
CC reducing the hematopoietic progenitor cell-depleting activity of  
CC a therapeutic treatment, including radiotherapeutic and/or  
CC chemotherapeutic treatments. Administration of FRIL compositions to a  
CC patient prior to treatment of the patient with a therapeutic treatment  
CC having a hematopoietic progenitor cell-depleting activity alleviates or  
CC reduces the hematopoietic progenitor cell-depleting activity of the  
CC therapeutic treatment in the patient. FRIL family members are useful for  
CC isolating population of progenitor cells, hemangioblasts, and mesenchymal  
CC stem cells. The composition is administered to reduce progenitor cell  
CC depleting effects of chemotherapy, so that the patient can receive  
CC a higher dose of the chemotherapeutic and preferably recover from cancer.  
CC It is also administered to patients having, or predisposed to developing  
CC a condition where the patients hematopoietic progenitor cells are

CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
CC The isolated mesenchymal cells are useful for tissue repair.  
XX SQ Sequence 303 AA;  
Query Match 58.2%; Score 797; DB 22; Length 303;  
Best Local Similarity 60.5%; Pred. No. 4e-69;  
Matches 161; Conservative  
QY 1 AGSLSFSTKFDPNQEDLIFOGHATSTNNVLOVTKLDSAGNPVSSAGRVLYSAPLRLWE 60  
DQ 1 AQSLSFSTKFDSDQKDLMFQGHATSTSSNIOVTKLDSAGNPVSSAGRVLYSAPLRLWE 60  
QY 61 DSAVLTSFDTTINFEISPTYSRIADGLAFFIAPPDSVI--SYHGGFLGLFPNANTLNS 118  
DQ 61 NSMAVSFEINLTQISTPHEPYAAGDGAFFLAPHDVTIPNSWGKFLGLY--SNVFRNS 118  
QY 119 STSENQTTTAAASNNVAVFEDTYLNDYDGNPIHIGIDVNSIRSKVTAKWDQNGKI 177  
DQ 119 PTSENQSGDVNTDSRVAVFEDTFPPNANI-DPNYRHIGIDVNSIKSKETARWONGKT 177  
QY 178 ATAHISYNSVSKRLSVTSYYAGSKPATLSYDIELHTVLPWVVRVGLSASTGQDKRNTVH 237  
DQ 178 ATAHISYNSASRLSVTTPYGGKAVSLSHDVELTQVLPWVVRVGLSASTGGEKQKNTII 237  
QY 238 SWSFTSSLTWNVAKKENENKYITRGV 263  
DQ 238 SWSFTSSLKLNNEVKEPKEDMYIANVV 263  
RESULT 6  
AAY58736 standard; Protein; 290 AA.  
XX ID AAY58736;  
XX AC AAY58736;  
XX DT 25-APR-2000 (first entry)  
XX DE Jack bean concanavalin A lectin.  
XX KW Concanavalin A; Con A; lectin; jack bean; insecticide;  
KW transgenic plant; Brassica; insect resistance.  
XX OS Canavalia ensiformis.  
XX FH Key Location/Qualifiers  
XX FT Modified-site 152  
XX FT /note= "N-glycosylated"  
XX PN WO200001223-A1.  
XX PD 13-JAN-2000.  
XX PF 02-JUL-1999; 99WO-SE01209.  
XX PR 03-JUL-1998; 98SE-0002425.  
XX (PLAN-) PLANT SCI SVERIGE AB.  
XX Ahman I, Melander M, Vamling K;  
XX WPI: 2000-160693/14.  
XX N-PSDB; AAZ58017.  
XX Novel lectins used to produce transgenic Brassica plants which are  
PT resistant to insects  
XX Example 2a; Fig 4; 51pp; English.  
XX The present sequence is that of the concanavalin A (Con A) lectin  
CC of jack bean. The present invention relates to: a transformed  
CC Brassica plant that is resistant to certain insect pests; an  
CC expression cassette containing DNA that codes for at least 1



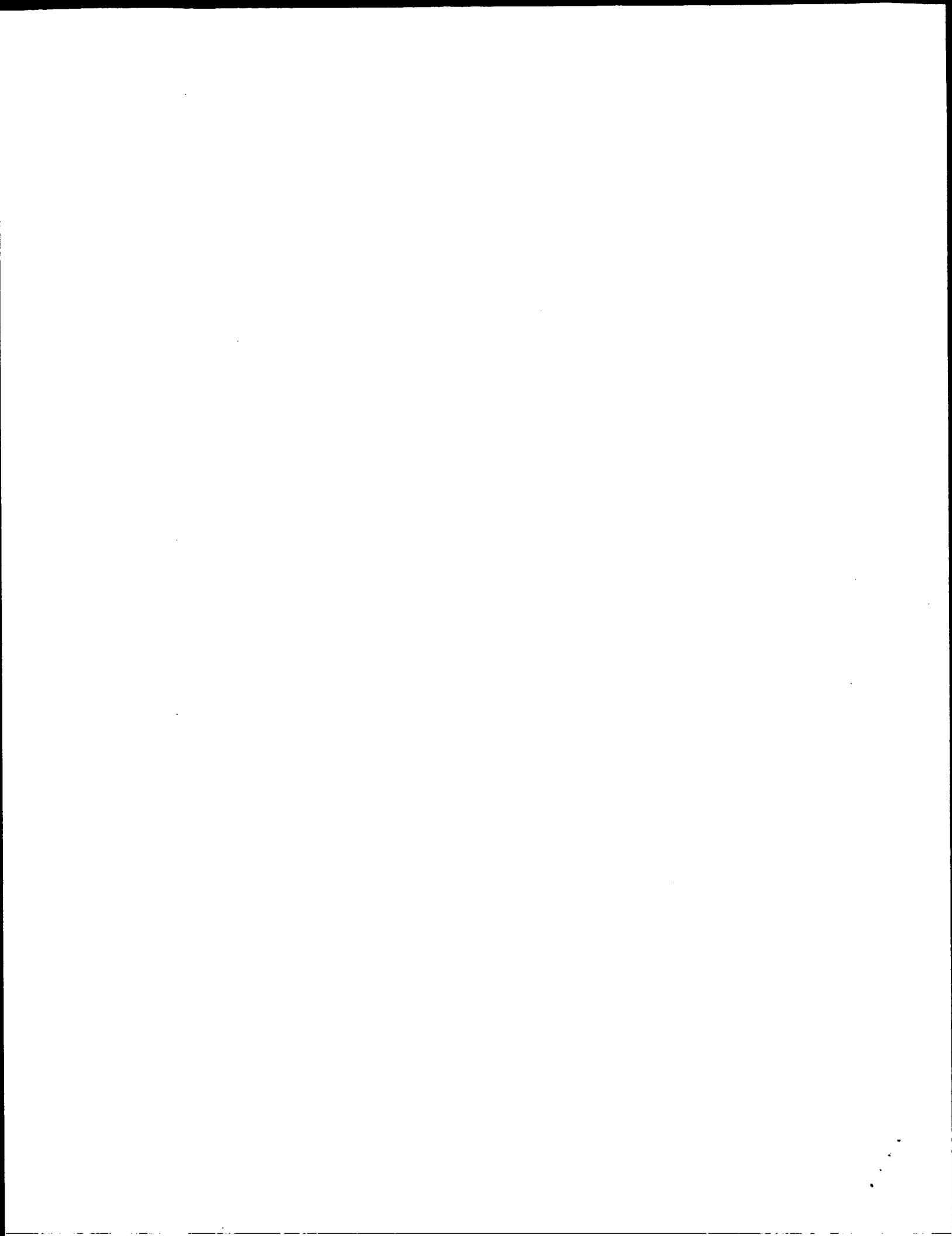


DR	XX	WPI; 1995-098720/13.	XX	28-AUG-1998;	98WO-GB02612.
XX	XX	New cpds. inhibiting binding of serum amyloid P to amyloid	XX	29-AUG-1997;	97GB-0018413.
PT	PT	fibrils - produced by computer assisted molecular design, useful	XX	(ALIZ-) ALIZYME THERAPEUTICS LTD.	
PT	PT	for preventing, treating or diagnosing amyloidosis, e.g.	XX	Bardocz 2M, Fish NW, Palmer RMJ, Pusztai AJ;	
PT	PT	Alzheimer's disease	XX	WPI; 1999-228935/19.	
XX	XX	Claim 7; Fig 4; 72pp; English.	XX	Use of Robina pseudoacacia lectin in medicine	
XX	XX	The sequences given in AAR74763-70 represent various pentraxins and	XX	Disclosure; Fig 1; 37pp; English.	
CC	CC	legume lectins. All these proteins show structural homology and may	XX	The invention relates to Robinia pseudoacacia lectins which can be used	
CC	CC	be used in the molecular design of a molecule for the inhibition of	CC	in medicine. The lectin is used for the control of mucosal cell	
CC	CC	serum amyloid P (SAP) binding to amyloid fibrils. The similarities	CC	proliferation, for the reduction and/or treatment of damage caused by	
CC	CC	in the amino acid sequences of SAP, human and Limulus C-reactive	CC	cell damaging agents, especially in the treatment of cancer, and/or for	
CC	CC	protein (CRP) and female hamster protein suggests that they may have	CC	the reduction and/or treatment of metabolic disorders. It is especially	
CC	CC	similar 3-D structures. The jellyroll topology of the pentraxins is	CC	useful for the treatment of mucositis in mammalian cells and/or tissues,	
CC	CC	reminiscent of the picornavirus coat proteins which also have	CC	particularly human cells and/or tissues (especially mucous cells	
CC	CC	pentameric structures. However, pentaxins resemble more closely	CC	including mucous membrane). Use of the lectin is effective and does not	
CC	CC	legume lectins such as concanavaline A and pea lectin. Alignment	CC	damage the gut as is the case with cytotoxic drugs and radiation. The	
CC	CC	of sequences on the basis of topologically equivalent features of the	CC	present sequence represents a polypeptide subunit of R. pseudoacacia	
CC	CC	three dimensional structures shows that helices occupy different	CC	lectin RPBAl.	
CC	CC	positions in the pentraxins and legume lectins and that the amino acid	XX	Sequence 285 AA;	
CC	CC	sequences of the two families have identities of only approx. 11%. The	XX	Query Match 33.8%; Score 463; DB 20; Length 285;	
CC	CC	two main helices in SAP occur before and after strand L, whereas the	XX	Best Local Similarity 43.9%; Pred. No. 1.2e-36;	
CC	CC	helices in the legume lectins occur at the C-terminus of strand J.	XX	Matches 115; Conservative 40; Mismatches 81; Indels 26; Gaps 10;	
CC	CC	There is a long insertion between the end of the helix after strand D	XX		
CC	CC	and the beginning of strand E in the lectins relative to pentraxins.	XX		
CC	CC	Strands G, H and I together with the type IV beta-hairpin between H and	XX		
CC	CC	I are identical in both SAP and pea lectin. The so-called pentraxin	XX		
CC	CC	octapeptide signature sequence, HXCXS/TWXS, is in this region so it is	XX		
CC	CC	not conserved in the legume lectins.	XX		
XX	XX	Sequence 237 AA;	XX		
XX	XX	Query Match 40.3%; Score 552; DB 16; Length 237;	XX		
XX	XX	Best Local Similarity 50.9%; Pred. No. 2e-45;	XX		
XX	XX	Matches 115; Conservative 30; Mismatches 61; Indels 20; Gaps 3;	XX		
QY	25	TSTNNLVQVTKLDSAGNPVSSAGRLVYSAPLRLEDVSAVLTSTFTIINFEISTPYTSRI 84	QY	2	GSLSEFSTKEDPNQEDLIFQCHATSTN-NVLQVTKLDSAGNPVSSAGRLVYSAPLRLEW 60
Db	2	TGTDGNLELTVSSNGSPGSGVGRALFAPVPHIWESSAATVSEFATFAFLIKSP-DSHP 60	Db	33	GSLSEFSPKFKHQPDLIFQSDALVTSGVLQTLTVND--GRVYDSIGRLVYAAPFQIWD 90
QY	85	ADGLAFFIAPDPSVI--SYHGFGFLGFPNANTLNNSSTSENQTTTKAASNNVAVFEDHY 142	QY	61	D-SAVLTSTFTIINFEISTPYTSRIADGLAFFIAPDPSVISYH-GGFLGLFPNANTLNNS 118
Db	61	ADGLAFFISNIDSSIPSGSTGRILGLFPDAN-----ADTIVAVELDTY 103	Db	91	STTGNAVASFVTSFSEFIKAPNEGKTADGLVFLAPVSTQPLKGGLLGLF----- 141
QY	143	LNPDYGDPNYTHIGIDVNSIRSKVTAKWDQNGKIATATAHISYNSVSKRLSVTSYYAGSKP 202	QY	119	STGENSTQTTKAASNNVAVFEDTYINPDYGDPNYTHIGIDVNSIRSKVTAKWDQNGKIA 178
Db	104	PNTDGDPSYPHIGIDIKSVRSKKTAKWNQDGKVGCTAHLIYNSVDKRLSAVSYPNADA 163	Db	142	--KDESYNK--SNQIYAVEFEDTFRNVAW-DPNGIHMGIDVNSIQSVRTVRWDWANGEVA 195
QY	203	ATLSYDIELHTVLPWVRVGLSASTGQDKERNVTWSWSTSLWLN 248	QY	179	TAHISYNSVSKRLSVTSYYAG-SKPATLSYDIELHTVLPWVRVGLSASTGQDK---ERN 234
Db	164	TSVSYDVLNDVLPWVRVGLSASTGLYKETNTIISWSFTSKLSN 209	Db	196	NVFISYEASTKSLTASLVPSLEKSFILSAIVDLKVLPEWVRVGFATTGLSEDYVQTN 255
RESULT 9			QY	235	TVHSWSFTSSL--WTNVAKKEN 254
AY06812			Db	256	DVLSWSFESNLPGGSNASVSKN 277,
XX	XX	AA06812 standard; protein; 285 AA.	XX		
XX	XX	AA06812;	XX		
XX	XX	23-JUN-1999 (first entry)	XX		
XX	XX	R. pseudoacacia lectin RPBAl polypeptide subunit b.	XX		
XX	XX	Robinia pseudoacacia; lectin; medicine; mucosal cell proliferation;	XX		
XX	XX	cancer; metabolic disorder; mucositis; cytotoxic.	XX		
XX	XX	Robinia pseudoacacia.	XX		
XX	XX	WO9911278-A1.	XX		
XX	XX	11-MAR-1999.	XX		
PD			XX		
RESULT 10			XX		
AY06814			XX		
ID			XX		
XX	XX	AA06814 standard; protein; 285 AA.	XX		
XX	XX	AA06814;	XX		
XX	XX	23-JUN-1999 (first entry)	XX		
XX	XX	Amino acid sequence of lectin RPBAl from R. pseudoacacia seed.	XX		
XX	XX	Robinia pseudoacacia; lectin; medicine; mucosal cell proliferation;	XX		
XX	XX	cancer; metabolic disorder; mucositis; cytotoxic.	XX		
XX	XX	Robinia pseudoacacia.	XX		
XX	XX	WO9911278-A1.	XX		
XX	XX	11-MAR-1999.	XX		
PD			XX		









GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:43:23 : Search time 25.0487 Seconds  
(without alignments)  
2171.628 Million cell updates/sec

Title: US-09-476-485A-2

Perfect score: 1370

Sequence: 1 AGSLSFSTKFDPNQEDLIF.....LWTVAKKENKIITRGVL 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum Match 0%

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1362	99.4	272	10 Q9ZTA9	Q9ZTA9 dolichos la
2	797	58.2	279	10 Q9M7M4	Q9M7M4 phaseolus v
3	642	46.9	290	10 Q947H0	Q947H0 canavalia e
4	641	46.8	290	10 O04672	O04672 canavalia b
5	631.5	46.1	293	10 P93537	P93537 sophora jap
6	583.5	42.6	284	10 Q9FVU9	Q9FVU9 sophora fla
7	576.5	42.1	266	10 P93536	P93536 sophora jap
8	502	36.6	286	10 P93248	P93248 maackia amu
9	496	36.2	254	10 Q43376	Q43376 arachis hyp
10	494	36.1	282	10 P93247	P93247 maackia amu
11	493	36.0	280	10 Q43374	Q43374 arachis hyp
12	486	35.5	254	10 Q43377	Q43377 arachis hyp
13	459.5	33.5	258	10 Q9FVF8	Q9FVF8 ulex europe
14	459	33.5	256	10 P93246	P93246 maackia amu
15	454.5	33.2	273	10 Q8RVY4	Q8RVY4 phaseolus c
16	452	33.0	285	10 Q9ZWP6	Q9ZWP6 robinia pse

17	443.5	32.4	247	10 Q8WLR6	Q8WLR6 griffonia s
18	436	31.8	251	10 Q93X49	Q93X49 lens culina
19	436	31.8	251	10 Q8VXF2	Q8VXF2 lens culina
20	435	31.8	251	10 Q93X50	Q93X50 lens culina
21	434	31.7	251	10 Q93X41	Q93X41 lens odemen
22	434	31.7	251	10 Q93WH6	Q93WH6 lens culina
23	434	31.7	251	10 Q8W4Y4	Q8W4Y4 lens lamott
24	433.5	31.6	273	10 Q8RVH2	Q8RVH2 phaseolus v
25	432.5	31.6	275	10 Q8RVH1	Q8RVH1 phaseolus v
26	432	31.5	279	10 Q49899	Q49899 medicago sa
27	431.5	31.5	251	10 Q8WLR7	Q8WLR7 griffonia s
28	431.5	31.5	275	10 Q8RVX5	Q8RVX5 phaseolus v
29	430.5	31.4	275	10 Q8RW23	Q8RW23 phaseolus c
30	425	31.0	251	10 Q93X48	Q93X48 lens ervoid
31	425	31.0	274	10 Q43628	Q43628 phaseolus v
32	423.5	30.9	268	10 P93458	P93458 phaseolus l
33	423.5	30.9	278	10 Q9LEB8	Q9LEB8 phaseolus l
34	420.5	30.7	275	10 Q8RVY1	Q8RVY1 phaseolus v
35	420.5	30.7	276	10 Q40750	Q40750 phaseolus a
36	420	30.7	251	10 Q93X42	Q93X42 lens nigric
37	419.5	30.6	278	10 Q9LEP9	Q9LEP9 phaseolus l
38	416.5	30.4	255	10 Q40912	Q40912 phaseolus l
39	415.5	30.3	273	10 Q8RVX9	Q8RVX9 phaseolus v
40	415	30.3	251	10 Q8RW33	Q8RW33 lathyrus sa
41	412.5	30.1	275	10 Q8RVX6	Q8RVX6 phaseolus v
42	411.5	30.0	273	10 Q42411	Q42411 medicago sa
43	411.5	30.0	275	10 Q8RVH3	Q8RVH3 phaseolus v
44	410.5	30.0	275	10 Q8RVV0	Q8RVV0 phaseolus v
45	409	29.9	285	10 Q9ZWP5	Q9ZWP5 robinia pse

## ALIGNMENTS

### RESULT 1

Q9ZTA9 PRELIMINARY; PRT; 272 AA.  
AC Q9ZTA9  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Mannose lectin.  
GN FRIL.  
OS Dolichos lab lab (Field bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Dolichos.  
OX NCBI\_TaxID=35936;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=COTYLEDON;  
RX MEDLINE=99110944; PubMed=9892687;  
RA Colucci G., Moore J.G., Feldman M., Chrispeels M.J.;  
RT "CDNA cloning of FRIL, a lectin from Dolichos lablab, that preserves  
RT hemopoietic progenitors in suspension culture";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:646-650(1999).  
DR HSSL; AF067417; AAD10734.1;  
DR HSSL; P02866; IONA  
DR InterPro: IPR000985; Lectin\_legA.  
DR InterPro: IPR001220; Lectin\_legB.  
DR Pfam: PF00138; lectin\_legA; 1.  
DR Pfam: PF00139; lectin\_legB; 1.  
DR ProDom: PD000671; Lectin\_legA; 1.  
DR ProDom: PD000711; Lectin\_legB; 1.  
DR PROSITE: PS00308; LECTIN\_LEGME\_BETA; 1.  
DR PROSITE: PS00307; LECTIN\_LEGME\_ALPHA; 1.  
KW Lectin.  
SQ SEQUENCE 272 AA; 29900 MW; EA6C004307441495 CRC64;

Query Match 99.4%; Score 1362; DB 10; Length 272;  
Best Local Similarity 99.6%; Pred. No. 3.4e-96;  
Matches 263; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGSLSFSTKFDPNQEDLIFQGHATSTNNVLQVTKLDSAGNPVSSSAGRVLYSAPLRWE 60  
DB 9 AQSLSFSTKFDPNQEDLIFQGHATSTNNVLQVTKLDSAGNPVSSSAGRVLYSAPLRWE 68  
QY 61 DSAVLTSFTDIINFEISTPTYSRIADGLAFFTAPDPSVTSYHGGLFLFPNANTLNNSST 120  
DB 69 DSAVLTSFTDIINFEISTPTYSRIADGLAFFTAPDPSVTSYHGGLFLFPNANTLNNSST 128  
QY 121 SENQTTTAASSNVAVAEFTYLNPDYDGPNYTHIGIDVNSIRSKVTAKWDQNGKIATA 180  
DB 129 SENQTTTAASSNVAVAEFTYLNPDYDGPNYTHIGIDVNSIRSKVTAKWDQNGKIATA 188  
QY 181 HTSYNSVSKRLSVTSYAGSKPATLSYDIELHTVLPBWVGVLSASTGQDKERTVHSWS 240  
DB 189 HTSYNSVSKRLSVTSYAGSKPATLSYDIELHTVLPBWVGVLSASTGQDKERTVHSWS 248  
QY 241 FTSSLTWNVAKKENNKYITRGV 264  
DB 249 FTSSLTWNVAKKENNKYITRGV 272  
RESULT 3  
Q9M7M4 PRELIMINARY; PRT; 279 AA.  
AC Q9M7M4;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Mannose lectin FRIL (Fragment).  
OS Phaseolus vulgaris (Kidney bean) (French bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.  
OX NCBI\_TaxID=3885;  
RN SEQUENCE FROM N.A.  
RP Moore J.G., Colucci G., Fuchs C.A., Hicklin D.J., Chrispeels M.J.,  
RA Feldman M.;  
RT "A new lectin in red kidney bean called PvPRIL stimulates  
RT proliferation of NIH3T3 cells expressing the Flt3 receptor.";  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF121458; AAF28739.1; -;  
DR HSP; P02866; IONA.  
DR InterPro; IPR000985; Lectin\_legA.  
DR InterPro; IPR001220; Lectin\_legB.  
DR Pfam; PF00138; lectin\_legA; 1.  
DR Pfam; PF00139; lectin\_legB; 1.  
DR ProDom; PD000671; lectin\_legA; 1.  
DR ProDom; PD000711; lectin\_legB; 1.  
DR PROSITE; PS00308; LECTIN\_LEGUME\_ALPHA; 1.  
DR PROSITE; PS00307; LECTIN\_LEGUME\_BETA; UNKNOWN\_1.  
KW Lectin.  
FT NON\_TER  
SQ SEQUENCE 279 AA; 31102 MW; F8919CF8B3EE4652 CRC64;  
Query Match 58.2%; Score 797; DB 10; Length 279;  
Best Local Similarity 60.5%; Pred. No. 4.4e-53;  
Matches 161; Conservative 39; Mismatches 60; Indels 6; Gaps 4;  
QY 1 AGSLSFSTKFDPNQEDLIFQGHATSTNNVLQVTKLDSAGNPVSSSAGRVLYSAPLRWE 60  
DB 1 AQSLSFSTKFDPNQEDLIFQGHATSTNNVLQVTKLDSAGNPVSSSAGRVLYSAPLRWE 60  
QY 61 DSAVLTSFTDIINFEISTPTYSRIADGLAFFTAPDPSVTSYHGGLFLFPNANTLNNS 118  
DB 61 NSMAVSSPETNLTIQISTPHPYAAGDEAFFLPHDPTVPPNSWGKFLGLY--SNVFRNS 118  
QY 119 STSENQTTTAA--ASSNVAVAEFTYLNPDYDGPNYTHIGIDVNSIRSKVTAKWDQNGKI 177  
DB 119 PTSENOFGDVTNDTSKVAEFTDFTFNANI--DPNYRHIGIDVNSIRSKVTAKWDQNGKI 177  
QY 178 ATAHISYNSVSKRLSVTSYAGSKPATLSYDIELHTVLPBWVGVLSASTGQDKERTVH 237  
DB 178 ATAHISYNSVSKRLSVTSYAGSKPATLSYDIELHTVLPBWVGVLSASTGQDKERTVH 237

DB 178 ATAHISYNSKSKSTVTTFYPGMEVVALSHDVLHAEPLPEWVRVGLSASTGEEKOKNTII 237  
QY 238 SWSFTSSLTWNVAKKENNKYITRGV 263  
DB 238 SWSFTSSLTWNVAKKENNKYITRGV 263  
RESULT 3  
Q947H0 PRELIMINARY; PRT; 290 AA.  
AC Q947H0;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Concanavalin A.  
GN CONA.  
OS Canavalia ensiformis (Jack bean) (Horse bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Canavalia.  
OX NCBI\_TaxID=3823;  
RN SEQUENCE FROM N.A.  
RP STRAIN=CV. U-02;  
RA Ramis C., Gomord V.;  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF308777; AAL09432.1; -;  
DR InterPro; IPR000985; Lectin\_legA.  
DR InterPro; IPR001220; Lectin\_legB.  
DR Pfam; PF00138; lectin\_legA; 1.  
DR Pfam; PF00139; lectin\_legB; 1.  
DR ProDom; PD000671; Lectin\_legA; 1.  
DR ProDom; PD000711; Lectin\_legB; 1.  
DR PROSITE; PS00308; LECTIN\_LEGUME\_ALPHA; UNKNOWN\_1.  
DR PROSITE; PS00307; LECTIN\_LEGUME\_BETA; UNKNOWN\_1.  
SQ SEQUENCE 290 AA; 31480 MW; 0F2FDBCF547E42 CRC64;  
Query Match 46.9%; Score 642; DB 10; Length 290;  
Best Local Similarity 53.8%; Pred. No. 3e-41;  
Matches 134; Conservative 35; Mismatches 74; Indels 6; Gaps 4;  
QY 3 SLGSFSTKFDPNQEDLIFQGHATS--TNNVLQVTKLDSAGNPVSSSAGRVLYSAPLRWE 61  
DB 36 ALHFMENQFSKQDKLILQGDATGTGDNLELTRVSSNGSPQGSVGRALFYAPVHWES 95  
QY 62 SAVLTSFTDIINFEISTPTYSRIADGLAFFTAPDPSVTSYHGGLFLFPNANTLNNS 119  
DB 96 SAVVASFEATFTFLIKSP--DSHPADGIAFFLISNDSIPSGSTGRLGLFPDANVIRNST 154  
QY 120 TSENQTTTAASSNVAVAEFTYLNPDYDGPNYTHIGIDVNSIRSKVTAKWDQNGKIAT 179  
DB 155 TIDFNAAYNA--DTIVAVELDTVPNTDIDGDPSPHIGIDIKSVRSKKTAKWNQNGKVG 212  
QY 180 AHISYNSVSKRLSVTSYAGSKPATLSYDIELHTVLPBWVGVLSASTGQDKERTVHSW 239  
DB 213 AHIIYNSVSKRLSAVSYPNADSATVSDVDLDNVLPEWVRVGLSASTGLIKETNLTLSW 272  
QY 240 SFTSSLTWN 248  
DB 273 SFTSSLTWN 281  
RESULT 4  
Q04672 PRELIMINARY; PRT; 290 AA.  
AC Q04672;  
DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Lectin (Fragment).  
OS Canavalia brasiliensis (Brazilian jack bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;





OS Sophora flavescens.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Sophora.  
 OX NCBI\_TaxID=49840;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=ROOT;  
 RC TISSUE=LEAF;  
 RA Van B., Ma Z.G., Wang L.X., Chai H.M., Huang X.Q.;  
 RT "Cloning and sequencing of Sophora flavescens lectin gene, 284 AA."  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF285121; AAG00508.1;  
 DR HSSP: P02866; 1D02.  
 DR InterPro: IPR000985; Lectin\_legA.  
 DR InterPro: IPR001220; Lectin\_legB.  
 DR Pfam: PF00138; lectin\_legA; 1.  
 DR Pfam: PF00139; lectin\_legB; 1.  
 DR ProDom: PD000671; Lectin\_legA; 1.  
 DR ProDom: PD000711; Lectin\_legB; 1.  
 DR PROSITE: PS00308; LECTIN\_LEGUME\_ALPHA; 1.  
 DR PROSITE: PS00307; LECTIN\_LEGUME\_BETA; UNKNOWN\_1.  
 SQ SEQUENCE 284 AA; 31318 MW; 2DC947EB3CBE0F62 CRC64;  
 Query Match 42.6%; Score 583.5; DB 10; Length 284;  
 Best Local Similarity 49.1%; Pred. No. 8.6e-37;  
 Matches 133; Conservative 42; Mismatches 71; Indels 25; Gaps 10;  
 QY 1 AGSLSFSTKFDNQEDLIFQG--HATSTNNVLQVTKLDSAGNPVSSAGRVLSAPRLR 58  
 DB 31 ADSLSTFSDFDNGEDLLFGQDAHVTS--NNILQLTK--TSNGVPLQNTVGRALESTPIHL 88  
 QY 59 WEDSA-VLTSEFTDIINFEISTPTVTSRTADGLAFPIAPDSVI--SYHGGFGLFPNANTL 115  
 DB 89 WEKSTNRLSFSFTFTFLVISP--QSNPADGFAFIAPDPDTPPEGSOGGLGLFSPENAL 147  
 QY 116 NNSSTSENQTTTAASSNVVAVFDTYLPDYG--DPNIYIHIGIDVNSIRSKVTAKWDWQ 173  
 DB 148 N-----PKANQVAVVEFDFTYDKSSNSDNPVYHIGIDVQIKSSATVNRDK 195  
 QY 174 NGKIATAHISYNSVKRLS-VTSYAGSKATLSYDELHVLPEWVRVGLSASTGQDK 232  
 DB 196 EGVIGTARINYNATRLNSLVSSYPGSDQVSVYVVDLRTKLPEFRVGVFSASTGQOYQ 255  
 QY 233 RNTVHWSFTSSLTWNVAKKENNKYITR 263  
 DB 256 VHSIRSWFFSSSLHYVAKO--EDMYIARVV 284  
 RESULT 7  
 ID P93536 PRELIMINARY; PRT; 266 AA.  
 AC P93536;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Bark lectin II precursor (LECSJABMII) (B-SJA-II) (Fragment).  
 OS Sophora japonica (Japanese pagoda tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Sophora.  
 OX NCBI\_TaxID=3897;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BARK;  
 RC TISSUE=LEAF;  
 RX MEDLINE=97201486; PubMed=9049272;  
 RA van Damme E.J.M., Barre A., Rouge P., Peumans W.J.;  
 RT "Molecular cloning of the bark and seed lectins from the Japanese  
 RL pagoda tree (Sophora japonica).";  
 RT Plant Mol. Biol. 33:523-536 (1997).  
 RN [2]  
 RP SEQUENCE OF 17-40.  
 RC TISSUE=BARK;  
 RX MEDLINE=91131618; PubMed=1993686;

RA Ueno M., Ogawa H., Matsumoto I., Seno N.;  
 RT "A novel mannose-specific and sugar-specifically aggregatable lectin  
 from bark of the Japanese pagoda tree (sophora japonica).";  
 RL J. Biol. Chem. 266:3146-3153(1991).  
 CC J.- FUNCTION: MANNOSE/GLUCOSE-BINDING BARK LECTIN. DISPLAYS  
 CC HEMAGGLUTININATING ACTIVITY AT PH 5-10, WITH A PH OPTIMUM AT 8-9.  
 CC STOCKS OF NITROGEN ARE STORAGE PROTEIN THAT PROBABLY MAINTAINS  
 CC MOLECULES THAT CAN BIND THEIR OWN CARBOHYDRATE SIDE CHAINS. THEY  
 CC COULD ALSO PLAY A ROLE IN THE PLANT'S DEFENCE AGAINST PHYTOPHAGOUS  
 CC INVERTEBRATES OR HERBIVOROUS HIGHER ANIMALS.  
 CC J.- SUBUNIT: COMPOSED OF FOUR SUBUNITS OF ABOUT 30 KDA WHICH ARE POST-  
 CC TRANSLATIONALLY CLEAVED INTO A TWO SMALLER POLYPEPTIDES: A2 AND  
 CC B2.  
 CC J.- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.  
 DR EMBL: U63012; AAB51442.1;  
 DR HSSP: P02866; IONA.  
 DR InterPro: IPR000985; Lectin\_legA.  
 DR InterPro: IPR001220; Lectin\_legB.  
 DR Pfam: PF00138; lectin\_legA; 1.  
 DR Pfam: PF00139; lectin\_legB; 1.  
 DR ProDom: PD000711; Lectin\_legB; 1.  
 DR PROSITE: PS00308; LECTIN\_LEGUME\_ALPHA; 1.  
 DR PROSITE: PS00307; LECTIN\_LEGUME\_BETA; UNKNOWN\_1.  
 KW Lectin; Glycoprotein; Signal; Mannose-binding; Multigene family.  
 FT SIGNAL 1 16  
 FT CHAIN 17 129 BARK LECTIN II, A2 SUBUNIT.  
 FT CHAIN 130 266 BARK LECTIN II, B2 SUBUNIT.  
 FT CARBOHYD 129 129 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CARBOHYD 185 185 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CONFLICT 17 17 S -> A (IN REF. 2).  
 FT CONFLICT 25 25 N -> D (IN REF. 2).  
 FT CONFLICT 30 32 DQR -> NPE (IN REF. 2).  
 FT CONFLICT 38 38 G -> S (IN REF. 2).  
 SQ SEQUENCE 266 AA; 29509 MW; 78D72CAD9EF0919D CRC64;  
 Query Match 42.1%; Score 576.5; DB 10; Length 266;  
 Best Local Similarity 48.7%; Pred. No. 2.7e-36;  
 Matches 130; Conservative 38; Mismatches 70; Indels 29; Gaps 7;  
 QY 3 SLSPSFSTKFDNQEDLIFOGHATSTNNVLQVTKLDSAGNPVSSAGRVLSAPRLRLED- 61  
 DB 19 SLSTFNFNFGPDOROLLLOGDAHIPSGTLQTLTKTDSG-----VGRALYLPVHLWDSR 72  
 QY 62 SAVLTSFDTIINFEISTPTVTSRIADGLAFPIAPDSVI--SYHGGFGLFPNANTLNSS 119  
 DB 73 RGLASFETSFSEFVITSGTDDPGDGIAPFIAPPETTIPRSGSGGLGLFSPETALNSS- 131  
 QY 120 TSENQTTTAASSNVVAVFDTYLPDYGDPNVIHIGIDVNSIRSKVTAKWDWQNGKIAT 179  
 DB 132 -----LNPVAVVEFDFTFINEDW--DPSYWHIGIDVNSIKSSAAARWERKSGRKT 179  
 QY 180 AHISYNSVKRLSVTSYAGSK-----PATLSYDIHLHVLPEWVRVGLSASTGQDKERN 234  
 DB 180 AHISYNSKKSLSVSVSSPNTNCLVRVDYTVTSIDIDTLTVLPFWVRIGFSASTGYKIEH 239  
 QY 235 TVHWSFTSSLTWNVAKKENNKYITR 261  
 DB 240 SILSWFSFSSSQSSRAKK--EDLYIAR 264  
 RESULT 8  
 ID P93248 PRELIMINARY; PRT; 286 AA.  
 AC P93248;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Bark LEUCOAGGLUTININ II precursor (MALBII) (Fragment).  
 GN LECMALBII.  
 OS Maackia amurensis.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
GC eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Maackia.  
OX NCBI\_TaxID=37501;  
[1]  
RN SEQUENCE FROM N.A., AND SEQUENCE OF 29-48.  
RP TISSUE=BARK;  
RX MEDLINE=97390228; PubMed=9249142;  
RA van Damme E.J.M., Van Leuven F., Peumans W.J.;  
RT "Isolation, characterization and molecular cloning of the bark lectins  
from Maackia amurensis";  
RL Glycoconj. J. 14:449-456(1997).  
CC -1- FUNCTION: BARK LECTINS ARE STORAGE PROTEIN THAT PROBABLY MAINTAINS  
CC STOCKS OF NITROGEN DURING DORMANT PERIOD. SELF-AGGREGATABLE  
CC MOLECULES THAT CAN BIND THEIR OWN CARBOHYDRATE SIDE CHAINS. THEY  
CC COULD ALSO PLAY A ROLE IN THE PLANT'S DEFENCE AGAINST PHYTOPHAGOUS  
CC INVERTEBRATES OR HERBIVOROUS HIGHER ANIMALS (BY SIMILARITY).  
CC -1- SUBUNIT: HOMOTETRAMER OR HETEROTETRAMER OF MAHB AND MALB SUBUNITS.  
CC -1- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.  
DR EMBL: U65010; AAB39934.1; -;  
DR HSP: P19588; 1LUU.  
DR InterPro: IPR000985; Lectin\_legA.  
DR InterPro: IPR001220; Lectin\_legB.  
DR Pfam: PF00138; lectin\_legA; 1.  
DR Pfam: PF00139; lectin\_legB; 1.  
DR ProDom: PD000671; Lectin\_legA; 1.  
DR ProDom: PD000711; Lectin\_legB; 1.  
DR PROSITE: PS00308; LECTIN\_LEGUME\_ALPHA; 1.  
DR PROSITE: PS00307; LECTIN\_LEGUME\_BETA; UNKNOWN\_1.  
DR PROSITE: PS00430; TONE\_DEPENDENT\_REC\_1; UNKNOWN\_1.  
KW Lectin; Calcium; Manganese; Glycoprotein; Multigene family; Signal.  
FT NON\_TER 1 1  
FT SIGNAL <1 28  
FT CHAIN 29 286 BARK LEUCOAGGLUTININ II.  
FT CARBOHYD 27 27 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 207 207 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 219 219 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 286 AA; 31208 MW; 0F2E27617A0F6D00 CRC64;  
  
Query Match 36.6%; Score 502; DB 10; Length 286;  
Best Local Similarity 43.4%; Pred. No. 1.4e-30;  
Matches 116; Conservative 42; Mismatches 85; Indels 24; Gaps 9;  
  
QY 4 LSFSTKFPDQEDLIFOGHAT-STNNVLOVTKLSAGNPVSSAGRVLYSAPRLWED- 61  
DB 32 LSFSTKFPDQEDLIFOGHAT-STNNVLOVTKLSAGNPVSSAGRVLYSAPRLWED- 61  
QY 62 SAVLTSFDTIINFESTPTYSRIAAGLAFFIAPPDSVI--SYHGGFLGLFPNANTLNSS 119  
DB 91 TGSVASFTSTFFVYKAPNTITSDGLAFLAPPDSQIPSGRVSKYLGLFNSSN-----S 145  
QY 120 TSENQTTKAASSNVAVVEFDYLPNDYG--DPNYIHIGIDVNSIRSKYVAKWQNGKI 177  
DB 146 DSSNQ-----IVAVEFDYFSGHSDPDPNRIHIGIDVNGIESIKTVQWDWINGGV 196  
QY 178 ATAHISYNSVSKRLSVTSYAGSKA-TLSYDIEHTVLPWVRVGLSASTG--QDKERN 234  
DB 197 AFATITYLAPNKLIALSVYPSNQTSFIVAASVDLKEILLPEWVRVGSAAATGYPTQVETH 256  
QY 235 TVHWSFTSSLTWNVAKKENENKITYR 261  
DB 257 DVLWSFTSTLEANSDAENNVHAR 283  
  
RESULT 9  
Q43376 PRELIMINARY; PRT; 254 AA.  
AC Q43376;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)

DE Mannose/glucose-binding lectin precursor (Fragment).  
GN LEC.  
OS Arachis hypogaea (Peanut).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;  
OC Arachis.  
OX NCBI\_TaxID=3818;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. SELLIE; TISSUE=SEED;  
RA Law I.J.;  
RT "Cloning and expression of cDNA for mannose-binding lectin from  
peanut";  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.;  
DR EMBL: U22472; AAA74575.1; -;  
DR HSP: P02867; 2BQP.  
DR InterPro: IPR000985; Lectin\_legA.  
DR InterPro: IPR001220; Lectin\_legB.  
DR Pfam: PF00138; lectin\_legA; 1.  
DR Pfam: PF00139; lectin\_legB; 1.  
DR ProDom: PD000671; Lectin\_legA; 1.  
DR ProDom: PD000711; Lectin\_legB; 1.  
DR PROSITE: PS00307; LECTIN\_LEGUME\_BETA; UNKNOWN\_1.  
KW Lectin; Signal.  
FT NON\_TER 1 1  
FT SIGNAL <1 2  
FT CHAIN 254 AA; 28202 MW; 3AE1EALF90BICA03 CRC64;  
SQ SEQUENCE 254 AA; 28202 MW; 3AE1EALF90BICA03 CRC64;  
  
Query Match 36.2%; Score 496; DB 10; Length 254;  
Best Local Similarity 45.6%; Pred. No. 3.5e-30;  
Matches 114; Conservative 41; Mismatches 75; Indels 20; Gaps 7;  
  
QY 3 SLFSFTKFPDQEDLIFOGHAT-STNNVLOVTKLSAGNPVSSAGRVLYSAPRLWE 60  
DB 3 SLFSFTKFPDQEDLIFOGHAT-STNNVLOVTKLSAGNPVSSAGRVLYSAPRLWE 62  
QY 61 DSA-VLTSFDTIINFESTPTYSRIAAGLAFFIAPPDSVIVH--GGFLGLFPNANTLNN 117  
DB 63 KSTNRLTNFQAQFSFVVKSP-NDIGADGIAFFIAAPDSQIPKNSAGGTGLGF----- 113  
QY 118 SSTSENQTTKAASSNVAVVEFDYLPNDYG--DPNYIHIGIDVNSIRSKYVAKWQNG 175  
DB 114 ----DPQTAQNPSANQVLAFFEDTFYAGDSNGWDPNYQHIGIDVNSIRKSAATKWERRNG 169  
QY 176 KIATAHISYNSVSKRLSVTSYAGSKPATLSYDIEHTVLPWVRVGLSASTGQDKERN 235  
DB 170 QTLNVLTYDANSKNLQVTASYPGQRYQVSYVVDLHDLPEWVRVGSASSGQQVQSHE 229  
QY 236 VHSWSFTSSL 245  
DB 230 LQSWSTSNL 239  
  
RESULT 10  
P93247 PRELIMINARY; PRT; 282 AA.  
ID P93247  
AC P93247  
DT 01-MAY-1997 (Tremblrel. 03, Created)  
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
DE BARK LEUCOAGGLUTININ I precursor (MALBI) (Fragment).  
GN LECMALBI.  
OS Maackia amurensis.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Maackia.  
OX NCBI\_TaxID=37501;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BARK;  
RX MEDLINE=97390228; PubMed=9249142;  
RA van Damme E.J.M., Van Leuven F., Peumans W.J.;

RT "Isolation, characterization and molecular cloning of the bark lectins  
 RL from *Maackia amurensis*.";  
 CC Glycoconj. J. 14:449-456(1997).  
 CC -!- FUNCTION: BARK LECTINS ARE STORAGE PROTEIN THAT PROBABLY MAINTAINS  
 CC STOCKS OF NITROGEN DURING DORMANT PERIOD. SELF-AGGREGATABLE  
 CC MOLECULES THAT CAN BIND THEIR OWN CARBOHYDRATE SIDE CHAINS. THEY  
 CC COULD ALSO PLAY A ROLE IN THE PLANT'S DEFENCE AGAINST PHYTOPHAGOUS  
 CC INVERTEBRATES OR HERBIVOROUS HIGHER ANIMALS (BY SIMILARITY).  
 CC -!- SUBUNIT: HOMOTETRAMER OR HETEROTETRAMER OF MAHB AND MALB SUBUNITS.  
 CC -!- PTM: ONLY THREE OF THE PUTATIVE GLYCOSYLATION SITES ARE OCCUPIED.  
 CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.  
 DR EMBL: U65009; AAB39933.1;  
 DR HSSP: P19588; L1U1.  
 DR InterPro: IPR000985; Lectin\_legA.  
 DR InterPro: IPR001220; Lectin\_legB.  
 DR InterPro: IPR000531; TonB\_boxC.  
 DR Pfam: PF00138; lectin\_legA; 1.  
 DR Pfam: PF00139; lectin\_legB; 1.  
 DR ProDom: PD000671; Lectin\_legA; 1.  
 DR ProDom: PD000711; Lectin\_legB; 1.  
 DR ProSITE: PS00308; LECTIN\_LEGUME\_ALPHA; 1.  
 DR ProSITE: PS00307; LECTIN\_LEGUME\_BETA; UNKNOWN\_1.  
 DR ProSITE: PS00430; TONB\_DEPENDENT\_REC\_1; UNKNOWN\_1.  
 KW Lectin; Calcium; Manganese; Glycoprotein; Multigene family; Signal.  
 FT NON\_TER 1  
 FT SIGNAL <1 24 BY SIMILARITY.  
 FT CHAIN 25 282 BARK LEUCOAGGLUTININ I.  
 FT CARBOHYD 23 23 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 282 AA; 30879 MW; 45B3F714E1D8957A CRC64;  
 Query Match 36.1%; Score 494; DB 10; Length 282;  
 Best Local Similarity 43.1%; Pred. No. 5.7e-30;  
 Matches 115; Conservative 42; Mismatches 86; Indels 24; Gaps 8;  
 QY 4 LSFSTKFPDNEEDLIFOGHAT-STNNVLQVTKLDSAGNPVSSAGRVLSAPLRLWE- 61  
 DB 28 LSFSTNNFLNEADLLFOGEASVSTGVLQTLRVEN-GQPKYSVGRALYAAPVRIWNT 86  
 QY 62 SAVLTSFDTIINFEISTPTYSRTADGLAFIAPDPSVSYHG--GFLGLFPNANTLNNSS 119  
 DB 87 TGSVASFTSTFTVVKAPNPSTNSGLAFIAPDPSQIPTGSGVTKVGLGFNN----- 138  
 QY 120 TSENQTTTKAASNNVAVFEDTVLNDYG--DPNYTHIGIDVNSIRSKVTAKWDWONGKI 177  
 DB 139 -----TSDSSNQIVAVFEDTVLHKYNPDPNYRHIGIDVNGIDSIKTVQWDWINGGV 192  
 QY 178 ATAHISYNSVKRLSVTSYAGSKPA-TLSYDIELHTVLPWVRVGLSASTGQDKERNT- 235  
 DB 193 APATITLAPSKTLIASLYPSNQTSTFIVAASVDLKEILPEWVRVGFSAATGYPTVEVTH 252  
 QY 236 -VHSNFTSSLTWNVAKKENKYLTR 261  
 DB 253 DVLSWFTSTLEANSDAETNNVHIAR 279  
 RESULT 11  
 Q43374  
 ID Q43374 PRELIMINARY; PRT; 280 AA.  
 AC Q43374;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
 DE Mannose/glucose-binding lectin precursor.  
 GN LEC.  
 OS Arachis hypogaea (Peanut).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;  
 OC Arachis.

OX NCBI\_TaxID=3818;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. SELLIE; TISSUE=NODULE;  
 RA Law I.J.;  
 RT "Cloning and expression of cDNA for mannose-binding lectin from  
 peanut.";  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U22469; AAA74572.1; -.  
 DR HSSP: P02866; LDO2.  
 DR InterPro: IPR000985; Lectin\_legA.  
 DR InterPro: IPR001220; Lectin\_legB.  
 DR Pfam: PF00138; lectin\_legA; 1.  
 DR Pfam: PF00139; lectin\_legB; 1.  
 DR ProDom: PD000671; Lectin\_legA; 1.  
 DR ProDom: PD000711; Lectin\_legB; 1.  
 DR ProSITE: PS00307; LECTIN\_LEGUME\_BETA; UNKNOWN\_1.  
 KW Lectin; Signal.  
 FT SIGNAL 1 26  
 FT CHAIN 27 280  
 SQ SEQUENCE 280 AA; 31012 MW; C15B39B32F455BD5 CRC64;  
 Query Match 36.0%; Score 493; DB 10; Length 280;  
 Best Local Similarity 45.2%; Pred. No. 6.7e-30;  
 Matches 113; Conservative 44; Mismatches 73; Indels 20; Gaps 8;  
 QY 3 SLFSFTKFPDNOE-DLIFOGHAT-STNNVLQVTKLDSAGNPVSSAGRVLSAPLRLWE 60  
 DB 29 SLFSYNNFEQDDERNLILQGDATFSASKGIQTKVDDNGTFAKSTVGRVLSHTQVRLWE 88  
 QY 61 DSA-VLTSTFDTIINFEISTPTYSRTADGLAFIAPDPSVSYH--GGFLGLFPNANTLNN 117  
 DB 89 KSTNRLTNFQAQFSFVKSPIDNG-ADGIAPFAAPDSIPKNSAGGTGLGF-----D 140  
 QY 118 SSTSEQTTTKAASNNVAVFEDTVLNDYG--DPNYTHIGIDVNSIRSKVTAKWDWONG 175  
 DB 141 PSTAQN-----PSANQVLUAVFEDTVQADSGNDGWDPNYQHIGIDVNSIKSAATTKWERRNG 195  
 QY 176 KIATAHISYNSVKRLSVTSYAGSKPATLSYDIELHTVLPWVRVGLSASTGQDKERNT 235  
 DB 196 QTLNVLVSYDANSRNLQVTASYPDGQRYQVSYNVDLRYLPEWGSVGFSAASGQYQSHE 255  
 QY 236 VHSWSTSTSL 245  
 DB 256 LQSWSTSTSL 265  
 RESULT 12  
 Q43377  
 ID Q43377 PRELIMINARY; PRT; 254 AA.  
 AC Q43377;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
 DE Mannose/glucose-binding lectin precursor (Fragment).  
 GN LEC.  
 OS Arachis hypogaea (Peanut).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;  
 OC Arachis.  
 OX NCBI\_TaxID=3818;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. SELLIE; TISSUE=SEED;  
 RA Law I.J.;  
 RT "Cloning and expression of cDNA for mannose/glucose-binding lectin  
 from peanut.";  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U22473; AAA74576.1; -.  
 DR HSSP: P02867; 2BOP.  
 DR InterPro: IPR000985; Lectin\_legA.  
 DR InterPro: IPR001220; Lectin\_legB.

DR	Pfam:	PF00138;	lectin_legA; 1.
DR	Pfam:	PF00139;	lectin_legB; 1.
DR	ProDom:	PD000671;	Lectin_legA; 1.
DR	ProDom:	PD000711;	Lectin_legB; 1.
DR	PROSITE:	PS00307;	LECTIN_LEGUME_BETA; UNKNOWN_1.
KW	Lectin;	Signal.	
FT	NON_TER	1	
FT	SIGNAL	<1	
SEQ	SEQUENCE	254 AA;	28389 MW; 761B82817DE7DFB9 CRC64;
 Query Match            35.5%; Score 486; DB 10; Length 254;			
Best Local Similarity   44.4%; Pred. No. 2e-29;			
Matches	111;	Conservative	42; Mismatches   77; Indels   20; Gaps   7;
QY	3	SLSFSETKFPDNOE-DLIIFOGHAT-STNNVLQVTKLDSAGNPVSSSAGRVLVSAPLRWE	60
Dd	3	SLFSYNNFQDDERNLLGGDAKFSASKIGLTKVDDNGCTPAKSTVGRVLHSHQTVELWE	62
QY	61	DSA-VLTSPDTIINFELSTPYTSRIADGLAFFTAAPPDSVTSYH--GGFLGLEFNANTPLN	117
Dd	63	KSTNRILTFOAQFSFIKSPIDNG-ADGIAFFAAAPDSEIPKNSAGGTGLGF-----	113
QY	118	SSTSENQTTTKAASSNVAVVEFDTLNPDYG--DPNYIHIGIDVNSTRSKVTAKWDHQNG	175
Dd	114	----DPQATQANQAVLAVEFFDTFAQDSNGMDPNQHIGIDVNSIKSAITKKWERDG	169
QY	176	KIATAHSYNSVKRISVTSYVAGSKPATLSYDIELHTVLPWVRVGLSASTGDOKERNT	235
Dd	170	QTLNLVLYTDANSKNLOVTASYPDGQRQLSYRVLDRLYLPEWCRVGFSASGGQYOSHE	229
QY	236	VHSWSFTSSL 245	
Dd	230	LQSWSETSTL 239	
RESULT 13			
Q9FVF8	ID	Q9FVF8	PRELIMINARY; PRT; 258 AA.
AC	01-MAR-2001	(TrEMBLrel. 16, Created)	
DT	01-MAR-2001	(TrEMBLrel. 16, Last sequence update)	
DT	01-JUN-2002	(TrEMBLrel. 21, Last annotation update)	
DE	Lectin II (fragment).		
OS	Ulex europeus (furze).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
CC	SpERMaturaphyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OX	eurosid I; Fabales; Fabaceae; Papilionoideae; Genisteae; Ulex.		
NCBI_TaxId	3902;		
RN	[1]		
SEQUENCE FROM N.A.			
RA	Loris R., De Greve H., Dao-Thi M.-H., Messens J., Imberby A., Wyns L.;		
RP	"Structural basis of carbohydrate recognition by lectin II from Ulex		
RT	europeus, a protein with a promiscuous carbohydrate binding site.";		
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL:	AF190633; AAG16779.1; -	
DR	HSSP:	P05046; 1SBD.	
DR	InterPro:	IPR000985; Lectin_legA.	
DR	InterPro:	IPR001220; Lectin_legB.	
DR	Pfam:	PF00138; lectin_legA; 1.	
DR	Pfam:	PF00139; lectin_legB; 1.	
DR	ProDom:	PD000671; lectin_legA; 1.	
DR	ProDom:	PD000711; lectin_legB; 1.	
DR	PROSITE:	PS00308; LECTIN_LEGUME_ALPHA; 1.	
DR	PROSITE:	PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.	
FT	NON_TER	1	
SEQ	SEQUENCE	258 AA;	27909 MW; 581F6DD8F5E049FB CRC64;
 Query Match            33.5%; Score 459.5; DB 10; Length 258;			
Best Local Similarity   42.1%; Pred. No. 2.e-27;			
Matches	107;	Conservative	44; Mismatches   76; Indels   27; Gaps   10;
QY	4	LSFSETKFPDNOEDLIIFOGHAT-STNNVLQVTKLDSAGNPVSSSAGRVLVSAPLRWE	-61
Dd	4	LSFSETKFPDNOEDLIIFOGHAT-STNNVLQVTKLDSAGNPVSSSAGRVLVSAPLRWE	-61

Db 63 TGRVASESTSTFFVQK--AARLSDGLAFLAPDSDIPSGDVSKYLGLFNNSN----- 115  
QY 119 STSENQTTTAASSNVAVEDTYLNDYG--DPNYIHIGIDVNSIRSKVTAKWDQWQK 176  
Db 116 SESNQ-----IVAVEFDTFENHNDPDPNRYRHIGIDVNGIDSIKTVQWDYINGG 166  
QY 177 IATAHISVNSVKRLSVTSYAGSKPA-TLSYDIELHTVLPVWVRVGLSASTGQD--KER 233  
Db 167 VAFATITLAPNKTLIASLVYPSSETSFIVAASVDLKEILPEWVRVGFSAATGAPAAAE 226  
QY 234 NTVHSWSTSSLTNVAKKENKNIYTR 261  
Db 227 HDVRSWSTSTFEAN-SPVDHNVHIAR 253

## RESULT 15

Q8RVY4

ID Q8RVY4 PRELIMINARY; PRT; 273 AA.  
AC Q8RVY4;  
DT 01-JUN-2002 (TReMBLrel. 21, Created)  
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE Phytohemagglutinin precursor.  
GN LEC21.  
OS Phaseolus coccineus (Scarlet runner bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.  
OX NCBI\_TaxID=3886;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LEAF;  
RA Lioi L.;  
RT "Lectin genes in legumes.";  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ438774; CAD27654.1; -  
KW Signal.  
FT SIGNAL 1 21 POTENTIAL  
FT CHAIN 22 273 PHYTOHEMAGGLUTININ.  
SQ SEQUENCE 273 AA; 29522 MW; C2FEBACFA8037F6A CRC64;

Query Match 33.2%; Score 454.5; DB 10; Length 273;  
Best Local Similarity 43.4%; Pred. No. 5.6e-27;  
Matches 109; Conservative 33; Mismatches 84; Indels 25; Gaps 7;

QY 1 AGSLSFSTKFDNQEDLIPOGHAT-STNNVLQVTKLDSAGNPVSSAGRVLYSAPLRWL 59  
Db 22 ASSETSFSORF--NETNLIQGNASVSSGQLRLTNLKSNGEPTVGSIGRAFYSTPIQIW 79  
QY 60 ED-SAVLTSFDTLINFEISTPTYSRIADGLAFTAPDPSVISYHGGFLGLFPNANTLNS 118  
Db 80 DSTTGLASFATSTFNIVAPIVKNAGDLAFALVPVGSQPKGKGGFLGLFDGSGNT---- 135  
QY 119 STSENQTTTAASSNVAVEDTYLNDYGDPNKITHIGIDVNSIRSKVTAKWDQWQKIA 178  
Db 136 -----NFTVAVEFDLYNNWN-DPKERHIGIDVNSIKSIKTTTPWDFVNGENA 182  
QY 179 TATHISVNSVKRLSVTSYAGSKPA-TLSYDIELHTVLPVWVRVGLSASTGQD---ERN 234  
Db 183 KVHITYESSTKLMLASLVPSLKKSTVSDTVDLKSVLPWVSVGFSATTGIDKGNVETN 242  
QY 235 TVHSWSTSSSL 245  
Db 243 DILSWSFASKL 253

Search completed: February 26, 2003, 16:54:25  
Job time : 33.0487 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:40:33 ; Search time 19.7753 Seconds  
(without alignments)  
1283.395 Million cell updates/sec

Title: US-09-476-485A-2

Perfect score: 1370  
Sequence: 1 AGSLFSFTKFDPNQEDLIF.....LWTNVAKKENKYYITRGVL 264

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	657.5	48.0	293	2 S66356	mannose/glucose-binding lectin CLAI precursor - Cladrastis lutea
2	650	47.4	290	2 A34139	concanavalin A pre
3	640	46.7	290	1 CVJBP	concanavalin A pre
4	578	42.2	290	2 S66357	mannose/glucose-bi
5	575.5	42.0	132	2 A54864	mannose/glucose-sp
6	496	36.2	290	2 S66355	lectin-related sto
7	495	36.1	287	2 JC5444	leukoagglutinin pr
8	490	35.8	247	2 JC2268	hemagglutinin - Ma
9	478.5	34.9	286	2 S48033	lectin precursor -
10	473.5	34.6	244	2 S23099	lectin I, anti-H(O
11	463.5	33.8	285	2 S27365	lectin precursor
12	449	32.8	233	1 LNLWA	phytohemagglutinin
13	445.5	32.5	273	2 A5701	agglutinin I precu
14	441	32.2	285	2 S62690	lectin II - furze
15	440.5	32.0	249	2 JX0163	lectin - Scotch la
16	439	32.0	250	2 S16964	agglutinin II prec
17	439	32.0	285	2 S62691	probable lectin 2
18	432	31.5	279	2 T09620	mannose/glucose-sp
19	430.5	31.4	105	2 B4864	phytohemagglutinin
20	425	31.0	274	2 S1832	lectin II - Scotch
21	423.5	30.9	249	2 JQ1981	lectin precursor -
22	422	30.8	275	1 LNP	phytohemagglutinin
23	420.5	30.7	276	2 S1831	galactose-specific
24	420	30.7	241	2 JX0289	lectin precursor -
25	420	30.7	290	2 JX0175	lectin precursor
26	419.5	30.6	281	2 S09697	galactose-specific
27	419	30.6	241	2 JX0290	phytohemagglutinin
28	417.5	30.5	272	2 B22826	lectin II - lima b
29	416.5	30.4	253	2 A53416	

## ALIGNMENTS

### RESULT 1

S66356

mannose/glucose-binding lectin CLAI precursor - Cladrastis lutea

C:Species: Cladrastis lutea

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 20-Aug-1999

C:Accession: S66356; S66300

R:van Damme, E.J.M.; Barre, A.; Bemer, V.; Rouge, P.; van Leeuwen, F.; Peumans, W.J.

Plant Mol. Biol. 29, 579-598, 1995

A:Title: A lectin and a lectin-related protein are the two most prominent proteins in

A:Reference number: S66299; MUID:96123235; PMID:8534854

A:Accession: S66356

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-293 <VAM>

A:Cross-references: EMBL:U21958; NID:g1141756; PIDN:AAC49136.1; PID:g1141757

A:Experimental source: bark

A:Accession: S66300

A:Molecule type: protein

A:Residues: 36-46, S',48-55,'D',163-179,'X',181 <VAM>

C:Superfamily: plant lectin

C:Keywords: glycoprotein; lectin

F:1-35/Domain: signal sequence #status predicted <SIG>

F:36-293/Product: mannose/glucose-binding lectin CLAI #status experimental <MAT>

Query Match 48.0%; Score 657.5; DB 2; Length 293;

Best Local Similarity 53.2%; Pred. No. 5.3e-45;

Matches 140; Conservative 34; Mismatches 80; Indels 9; Gaps 6;

QY 3 SLFSFTKFDPNQEDLIFOGHAT-STNNVLOVTKLDSAGNPVSSSAGRVLYSAPLRWED 61

Db 38 SLSTFNNFPNSEDLLIFQKDAISSNETLETRISSGQPATSSVGRALYTPVRLWDK 97

QY 62 S-AVLTSPTDITINFEISPTYSRIADGLAFFIAPPDSVISHGGLFLFPNANTLNNST 120

Db 98 STGRLASFKTTFSAITSP-TQDPGDGFAFFIAPPDTPPGYGGLLGLF---NGFNLRNS 153

QY 121 SENQTTTAASSNVAVFEDYLNPDYGDPNVIHIGIDVNSRSTKVTAKDWQNGKIATA 180

Db 154 SNGVAVNQAQIVAVFEDTYINGQC-DPKYRHVGIDVNSITSLAYTQWQNGVKATA 212

QY 181 HISYNSVSKRLSVTSYAGSKPATLSYDIEHLTVLPWVRVGLSASTGQDKERTVHNS 240

Db 213 QISYNPASOKLTAVTSYNSPTPLTVSLDIDLOTVLPWVRVGFSGASTGONVERNSILAWS 272

QY 241 FTSSLTWNVAKENENKYYITRGV 263

Db 273 FSSSLTTLTAKK--EDMYIARYV 293

### RESULT 2

A34139

concanavalin A precursor - sword bean







Db 33 LSFTINFPVNEADLLFOGASVSSTGVLQLRVEN-GOPQKYSVGRALYAAPVRLWNT 91  
 Qy 62 SAVLTSFDTIINFELSTPYTSRIADGLAFFIAPPDSVI--SYHGGLGLFPNANTLNSS 119  
 Db 92 TGSVASFSTFTFVVKAPNPDIITSDGLAFYLAAPPDSQIPSGSVSKYLGLFNNSN-----S 146  
 Qy 120 TSENOTTTKAASNNAVVEFDYTLNPDYG--DPNVIHIGIDVNSIRSKVTAKWDWONGKI 177  
 Db 147 DSSNQ-----IVAVELDTYFAHSYDPWDPNTRHIGIDVNGIESIKTVOWDWINGGV 197  
 Qy 178 ATAHISYNSVKRLSVTSYYAGSKPA-TLSYDIELHTVLPWVRVGLSASTGQDKERN 235  
 Db 198 AFATITYLAPNKTLLASLYVPSNQTFSSVAASVDLKEILLPEWVRVGFSAATGYPTVEVETH 257  
 Qy 236 -VHSWTSFSSLTWNVAKKENKYITR 261  
 Db 258 DVLSWSTSTLEANCDAATENNVIHAR 284

## RESULT 8

JC2268  
 C:Species: Maackia amurensis  
 C:Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 07-May-1999  
 R:Accession: JC2268  
 R:Konami, Y.; Ishida, C.; Yamamoto, K.; Osawa, T.; Irimura, T.  
 J. Biochem. 115, 767-777, 1994  
 A:Title: A unique amino acid sequence involved in the putative carbohydrate-binding domain of hemagglutinin (HAH).  
 A:Reference number: JC2268; MUID:94375425; PMID:8089095  
 A:Accession: JC2268  
 A:Molecule type: protein  
 A:Residues: 1-247 <KON>  
 A:Experimental source: seed  
 C:Superfamily: plant lectin  
 C:Keywords: calcium binding; glycoprotein; manganese  
 F:111,177,189/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:125,127,138,143/Binding site: manganese (Glu, Asp, His) #status predicted  
 F:127,135,138/Binding site: calcium (Asp) #status predicted

Query Match 35.8%; Score 490; DB 2; Length 247;  
 Best Local Similarity 45.7%; Pred. No. 9.3e-32;  
 Matches 116; Conservative 34; Mismatches 78; Indels 26; Gaps 10;

Qy 4 LSFSPKTFDPNEDLIFOGHAT-STNNVLQVTKLDSAGNPVSSSAGRVLYSAPRLRWED- 61  
 Db 4 LSFTINFPNMGDILLFOGAVATVPTGVLQLHS-EENGQPLEYSVGRVLYTAPVRIWDS 62  
 Qy 62 SAVLTSFDTIINFELSTPYTSRIADGLAFFIAPPDSVI--SYHGGLGLFPNANTLNSS 119  
 Db 63 TGAVASFSTFTFVVK--ARGASDGLAFFLAPPDSQIPSGSVSKYLGLFNNSN-----S 115  
 Qy 120 TSENOTTTKAASNNAVVEFDYTLNPDYG--DPNVIHIGIDVNSIRSKVTAKWDWONGKI 177  
 Db 116 DSSNQ-----IVAVELDTYFHSYDPWDPNTRHIGIDVNGIESIKTVOWDWINGGV 166  
 Qy 178 ATAHISYNSVKRLSVTSYYAGSKPA-TLSYDIELHTVLPWVRVGLSASTGQDK--ERN 234  
 Db 167 AFATITYLAPNKTLLASLYVPSNQTFSSVAASVDLKEILLPEWVRVGFSAATGAPKAVETH 226  
 Qy 235 TVHSWTSFSSLTWN 248  
 Db 227 DVRSWSTSTLEAN 240

## RESULT 9

S48033  
 C:Species: Robinia pseudoacacia (black locust)  
 C:Species: Robinia pseudoacacia (black locust)  
 C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 20-Jun-2000  
 C:Accession: S48033; S68376  
 R:Yoshida, K.; Baba, K.; Yamamoto, N.; Tazaki, K.

plant Mol. Biol. 25, 845-853, 1994  
 A:Title: Cloning of a lectin cDNA and seasonal changes in levels of the lectin and its  
 A:Reference number: S48033; MUID:94355657; PMID:7915553  
 A:Accession: S48033  
 A:Molecule type: mRNA  
 A:Residues: 1-286 <YOS>  
 A:Cross-references: EMBL:DL17757; NID:9538528; PIDN:BAA04604.1; PID:9538529  
 A:Experimental source: inner bark  
 R:Tazaki, K.; Yoshida, K.; Shinohara, K.; Koshiba, T.; Yamamoto, N.  
 FEBS Lett. 377, 54-58, 1995

A:Title: Expression of cDNA for a bark lectin of Robinia in transgenic tobacco plants  
 A:Reference number: S68376; MUID:96130320; PMID:8543018  
 A:Accession: S68376  
 A:Molecule type: protein  
 A:Residues: 32-40 <TAZ>  
 A:Experimental source: inner bark  
 C:Superfamily: plant lectin  
 C:Keywords: glycoprotein; lectin  
 F:1-31/Domain: signal sequence #status predicted <SIG>  
 F:32-286/Product: lectin precursor #status experimental <MAT>

Query Match 34.9%; Score 478.5; DB 2; Length 286;  
 Best Local Similarity 44.7%; Pred. No. 9.4e-31;  
 Matches 117; Conservative 40; Mismatches 80; Indels 25; Gaps 10;

Qy 2 GLSFSFTKFPDNOEDLIFOGHATSTN-NVLQVTKLDSAGNPVSSSAGRVLYSAPRLWE 60  
 Db 33 GLSFSFEPKFKHSQPDLLFQSDALVTSKGVQLITVND-GRPVYDSIGRVLYAAPFQIWD 91  
 Qy 61 D-SAVLTSFDTIINFELSTPYTSRIADGLAFFIAPPDSVISYH-GGFLGLFPNANTLNSS 118  
 Db 92 STTGNAVASFVTSFSLIKAPNEGKTADGLVFFLAPVGSQPLKGGGLGLF----- 142  
 Qy 119 STSENOTTTKAASNNAVVEFDYTLNPDYDGNVTHIGIDVNSIRSKVTAKWDWONGKIA 178  
 Db 143 ---KDESINK--SNQIVAVEFDTFRNVAW-DPNGIHGIDVNSIQSVTRWDWANGVA 196  
 Qy 179 TAHISYNSVKRLSVTSYYAG-SKPATLSYDIELHTVLPWVRVGLSASTGQDK---ERN 234  
 Db 197 NVFISYEASTKSLTASLYVPSLEKSFILSAIVDLKVLPEWVRVGVFTATTGLSEDYVQTN 256  
 Qy 235 TVHSWTSFSSLT--WTNVAKKEN 254  
 Db 257 DVLSWSPESNLPGGSNASVAKN 278

## RESULT 10

S23099  
 C:Species: Cytisus sessilifolius  
 C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 22-May-1998  
 C:Accession: S23099; S13438  
 R:Konami, Y.; Yamamoto, K.; Osawa, T.; Irimura, T.  
 FEBS Lett. 304, 129-135, 1992  
 A:Title: Correlation between carbohydrate-binding specificity and amino acid sequence  
 A:Reference number: S23099; MUID:92316214; PMID:1618311  
 A:Accession: S23099  
 A:Molecule type: protein  
 A:Residues: 1-244 <KON>  
 A:Experimental source: seed  
 R:Konami, Y.; Yamamoto, K.; Osawa, T.  
 Biol. Chem. Hoppe-Seyler 372, 103-111, 1991  
 A:Title: Purification and characterization of two types of Cytisus sessilifolius anti  
 A:Reference number: S13438; MUID:91315748; PMID:1859626  
 A:Accession: S13438  
 A:Molecule type: protein  
 A:Residues: 1-3, 'K', 5-23, 'A', 25-29, 'K', 31-35 <BIO>  
 A:Experimental source: seed  
 C:Superfamily: plant lectin  
 C:Keywords: calcium binding; glycoprotein; homotetramer; lectin; manganese; seed  
 F:116-129/Domain: glycopeptide #status predicted <GLP>  
 F:113,117/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:127,129,140,145/Binding site: manganese (Glu, Asp, His) #status predicted

F;129,140/Binding site: calcium (Asp) #status predicted

```
Query Match      34.6%; Score 473.5; DB 2; Length 244;
Best Local Similarity 42.4%; Pred. No. 1.9e-30;
Matches 108; Conservative 43; Mismatches 81; Indels 23; Gaps
```

QY 4 LSFSTFKFDPNQEDLFIQGCHAT-STNNVLQVTKLDAGNPNVSSSAGRVLVSAPFLRLWED- 61  
 ||| :||| :||| :||| | | | | | | | | | |  
Db 5 LSFNDFKEVPNQNNILFQGEASVTGVLQVKV--SKPATRSIGRALYAAPVHLDNST 61  
  
QY 62 SAVLTSEDTLIINFEI-STPYTSRIADGLAFFIAPPDSVSYHGCFLLGFNPANTLNNSST 120  
 : :|||:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 62 TGRVASSETFSFPWKDEPKSGNGVDLGTLFFLPANSQP-SGSSAGLFGLFNSSDNKS- 119  
  
QY 121 SENOTTAKAASSNVVAVEFTDYLNPDYG-DPNTYIHIGIDVINSIRSVKTAKWDMWGNTIA 178  
 -----SNQLIAVEFTDYEKGTYNPWDPFKHIGVDVNISIKITKVMDWRNGEVA 169  
  
QY 179 TAHLSYSNSVKRSLTSDSYSVAGSPATL-SVDIELHTIVLPEWPVRVGLSASTGDOKERT-- 235  
 |:| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 170 NVVIYTRAPTCKSLTVSLTPSDOTSNIIVASVDLKAILPEWSVGFSGVGNAAEFETHD 229  
  
QY 236 VHSWFTSSLWTNWVA 250  
 ||| ||| :|||  
Db 230 VLSWFTSLNEANPA 244  
  
RESULT 11  
S27365 lectin precursor - soybean  
C:Species Glycine max (soybean)  
C:date 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 20-Aug-1999  
C:Accession S27365  
R:Vodkin, L.O.; Rhodes, P.R.; Goldberg, R.B.

```

QY 243 SSL 245
Db 225 SOL 227

RESULT 13
A25701
phytohemagglutinin chain L precursor - kidney bean
N:Alternate names: PHA-L
C:Species: Phaseolus vulgaris (kidney bean)
C:Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 11-Apr-1995
C:Accession: A25701
R:Voelker, T.A.; Staswick, P.; Chrispeels, M.J.
EMBO J. 5, 3075-3082, 1986
A:Title: Molecular analysis of two phytohemagglutinin genes and their expression in Phaseolus vulgaris
A:Reference number: A25701
A:Accession: A25701
A:Molecule type: DNA
A:Residues: 1-273 <VOE>
A:Experimental source: cv. Pinto UI111
C:Genetics:
A:Gene: dlec2
C:Superfamily: plant lectin

Query Match 32.5%; Score 445.5; DB 2; Length 273;
Best Local Similarity 42.6%; Pred. No. 3.7e-28;
Matches 107; Conservative 36; Mismatches 83; Indels 25; Gaps 7;

QY 1 AGSLFSFTKFPNQEDLIFQGHAT-STNNVLOVTKLDSAGNPVSSSAGRVLYSAPRLW 59
Db 22 ASQTFSSDFR--NETNLLQGDASVSSGQRLTNVNSGEPTVGLGRAFYSAPIQIW 79
QY 60 E-DSAVLTSFDTIINFESTPTYSRIADGLAFPIAPPDSVISVHGGLFLGFPNANTLNS 118
Db 80 DYTGNVASFDTNFTNVLNPNAGPADGLAFALVPVGSQPKDKGGLFLGFDGNS---- 134
QY 119 STSENQTTKAASSNVAVFDTYLPDGDNYIHGIDVNSIRSKVTAKWDQNGKIA 178
Db 135 -----SNFTVAVFDTYLPKNDW-DPRRHIGIDVNSIKSIKTPWDFVNGENA 182
QY 179 TAHISVNSKRLSVTSYVAGSKPA-TLSYVDIELHVLPEWVRVGLSASTGDK---ERN 234
Db 183 EVHIYESTKLLVASLYPSLKTSTVSDTVDLKSVLPWVSVGFSATGTGIKGNVETN 242
QY 235 TVHWSFTSSL 245
Db 243 DILWSFASKL 253

RESULT 14
S62690
agglutinin I precursor - black locust
N:Alternate names: lectin
C:Species: Robinia pseudoacacia (black locust)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999
C:Accession: S62690; S62885
R:van Damme, E.J.M.; Barre, A.; Rouge, P.; van Leuven, F.; Peumans, W.J.
Plant Mol. Biol. 29, 1197-1210, 1995
A:Title: The seed lectins of black locust (Robinia pseudoacacia) are encoded by two genes
A:Reference number: S62685; MUID:96191285; PMID:8616218
A:Accession: S62690
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-285 <VAN>
A:Cross-references: EMBL:U24249
A:Experimental source: seed
A:Accession: S62685
A:Molecule type: protein
A:Residues: 32-51 <VAV>
C:Superfamily: plant lectin
C:Keywords: glycoprotein; homotetramer; lectin
F:1-31/domain: signal sequence #status predicted <SIG>

```

```

F;32-285/Product: agglutinin I #status experimental <MAT>

Query Match 32.2%; Score 441; DB 2; Length 285;
Best Local Similarity 42.5%; Pred. No. 9e-28;
Matches 111; Conservative 42; Mismatches 84; Indels 24; Gaps 9;

QY 2 GSLSFSFTKFPNQEDLIFQGHATSTN-NVLQVTKLDSAGNPVSSSAGRVLYSAPRLWE 60
Db 33 GSLSFSFPKAPNPQYLIFQDALVLTSTGVQLTNVNV-GVPPRRSIGRALLYAAPFIWD 91
QY 61 D-SAVLTSFDTIINFESTPTYSRIADGLAFPIAPPDSVISVHGGLFLGFPNANTLNS 119
Db 92 NTTGNVASFDTNFTNVLNPNAGPADGLAFPIAPPDSVISVHGGLFLGFPNANTLNS 144
QY 120 TSENQTTKAASSNVAVFDTYLPDGDNYIHGIDVNSIRSKVTAKWDQNGKIA 179
Db 145 -SYNK-----SNQIVAVFDTNFIHDPKGRHMGINVASIVSKVTPNWTNGEVAN 196
QY 180 AHISVNSKRLSVTSYVAGSKPATLSYDI-ELHTVLPWVRVGLSASTGDK---ERN 235
Db 197 VFISYEASTKSLNASLVPSLETSTFIHAIVDKVLPEWVRVGFSAITGIDTGYVQND 256
QY 236 VHSWSETSSL--WTNVAKKEN 254
Db 257 VLSWSESNLPGGNSVASVKN 277

RESULT 15
JX0163
lectin II - furze
C:Species: Ulex europaeus (furze)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-May-1998
C:Accession: JX0163; S13436
R:Konami, Y.; Yamamoto, K.; Osawa, T.
J. Biochem. 109, 650-658, 1991
A:Title: The primary structures of two types of the Ulex europaeus seed lectin.
A:Reference number: JX0162; MUID:91331980; PMID:1869520
A:Accession: JX0163
A:Molecule type: protein
A:Residues: 1-249 <KON1>
R:Konami, Y.; Yamamoto, K.; Osawa, T.
Biol. Chem. Hoppe-Seyler 372, 95-102, 1991
A:Title: Purification and characterization of a new type lactose-binding Ulex europaeus seed lectin.
A:Reference number: S13435; MUID:91315756; PMID:1859632
A:Accession: S13436
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-25,27-36 <KON2>
C:Superfamily: plant lectin
C:Keywords: glycoprotein; manganese; seed
F:118,245/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:130,132,143,148/Binding site: manganese (Glu, Asp, Asp, His) #status predicted
F:132,143/Binding site: calcium (Asp) #status predicted

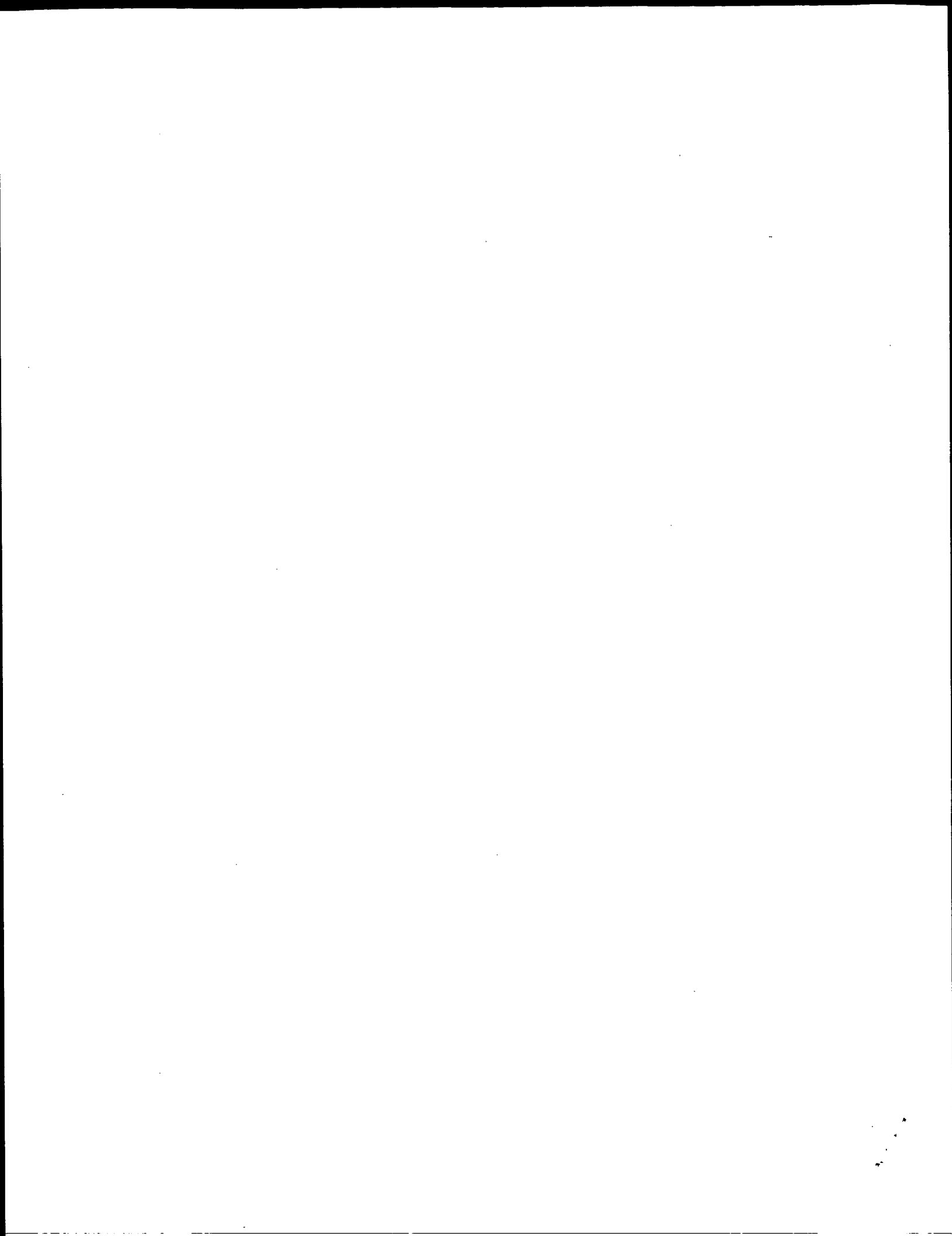
Query Match 32.2%; Score 440.5; DB 2; Length 249;
Best Local Similarity 41.4%; Pred. No. 8.2e-28;
Matches 106; Conservative 43; Mismatches 80; Indels 27; Gaps 10;

QY 4 LSFSTKFPNQEDLIFQGHAT-STNNVLOVTKLDSAGNPVSSSAGRVLYSAPRLWE 61
Db 6 LSFNFDKVPVQKNIIFQGDASVSTKGVLEVTKV---SKPTTRSIGRALLYAAPFIWD 62
QY 62 SAVLTSFDTIINFESTPTYSRI--ADGLAFPIAPPDSVI---SYHGGFLGLFPNANTL 117
Db 63 TGVKASFATSFVVKDEPDGKIDGVDFGLAFPLANSOIPSGSSAGMGLFCS----- 118
QY 118 SSTSENQTTKAASSNVAVFDTYLPDYG---DPNYIHGIDVNSIRSKVTAKWDQNG 175
Db 119 DSKSNQ-----IIAVFDSYFGKTYNPWDFDKHIGIDVNSIKSIKIVKDDWRNG 169
QY 176 KIATAHISVNSKRLSVTSYVAGSKPATLSYDI-ELHTVLPWVRVGLSASTGDK 232
Db 170 EVADVITYTRAPTKSLVSLVPSDGTSTNIVTASSVDLKAILPEWVRVGFSGVGNAAKF 229

```

QY 233 RNTVHWSFTSSLWTN 248  
: | | | | | | | | | |  
Db 230 DHDVLSHWFTTSNLEAN 245

Search completed: February 26, 2003, 16:51:31  
Job time : 20.7753 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:41:33 ; Search time 7.58052 Seconds  
(without alignments)  
1444.458 Million cell updates/sec

Title: US-09-476-485A-2

Perfect score: 1370

Sequence: 1 AGSLSFSTKFDPNQEDLIF.....LWTVAKKENKVIYTRGV 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	987	72.0	237	1	LECA_DOLLA
2	657.5	48.0	233	1	LECI_CLALU
3	650	47.4	290	1	CONA_CANGL
4	644.5	47.0	240	1	LEC_BOWMI
5	640	46.7	290	1	CONA_CANEN
6	581	42.4	290	1	LEC2_CLALU
7	502.5	36.7	240	1	LECS_VATMA
8	436	36.2	290	1	LECR_CLALU
9	431	35.8	292	1	LECS_SOPJA
10	430	35.8	270	1	LECB_SOPJA
11	478.5	34.9	286	1	LCB2_ROBPS
12	463.5	33.8	285	1	LEC_SOYBN
13	459.5	33.5	243	1	LECA_GRISI
14	448	32.7	280	1	LEC2_MEDTR
15	445.5	32.5	273	1	PHAM_PHAVU
16	441	32.2	285	1	LCSL_ROBPS
17	440.5	32.2	249	1	LEC2_ULEEU
18	440	32.1	285	1	LCB1_ROBPS
19	439	32.0	250	1	LECI_LABAL
20	439	32.0	285	1	LCB2_ROBPS
21	433	31.6	248	1	LEC2_CITSC
22	426.5	31.1	281	1	LEC2_CITSC
23	422.5	30.8	272	1	PHAL_PHAVU
24	422	30.8	275	1	LEC_PEA
25	420	30.7	290	1	LEC_BAUPU
26	412	30.1	233	1	LEC_VICFA
27	411	30.0	236	1	LEC_ONOVI
28	410.5	30.0	277	1	LECI_MEDTR
29	405	29.6	237	1	LECA_DIOGR
30	404	29.5	236	1	LECA_CRAFL
31	403	29.4	237	1	LECA_DIOGU
32	401.5	29.3	275	1	PHAE_PHAVU
33	401	29.3	237	1	CONA_CANLI

34	400	29.2	237	1	CONA_CANBR
35	400	29.2	237	1	CONA_CANVI
36	391.5	28.6	275	1	LEC1_DOLBI
37	391	28.5	244	1	LEC_LATSP
38	381	27.8	211	1	LEC_LENCU
39	368.5	26.9	275	1	LEC5_DOLBI
40	358.5	26.2	236	1	CONA_CANMR
41	357.5	26.1	242	1	LEC1_PSOTE
42	353	25.8	273	1	LECG_ARAHY
43	345.5	25.2	243	1	LEC1_ULEEU
44	338	24.7	243	1	LEC2_CYTSE
45	338	24.7	265	1	LECN_PEA

## ALIGNMENTS

RESULT 1  
LECA\_DOLLA  
ID P38662; STANDARD; PRT; 237 AA.  
AC P38662;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DE LECTIN.  
OS Dolichos lab lab (Field bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC Eubiosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Dolichos.  
OX NCBI\_TaxID=35936;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=cv. Lignosus; TISSUE=Seed;  
RX MEDLINE=94308133; PubMed=8034631;  
RA Gowda L.R., Savithri H.S., Rajagopal Rao D.;  
RT "The complete primary structure of a unique mannose/glucose-specific lectin from field bean (Dolichos lab lab).";  
RL J. Biol. Chem. 269:18789-18793(1994).  
CC -!- FUNCTION: D-MANNOSE/D-GLUCOSE-BINDING LECTIN. REQUIRES CA2+ AND MN2+ IONS FOR FULL ACTIVITY.  
CC -!- SUBUNIT: Tetramer of two alpha and two beta chains.  
CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.  
DR HSSP; P02866; 10NA.  
DR InterPro; IPR000985; Lectin\_legA.  
DR InterPro; IPR001220; Lectin\_legB.  
DR Pfam; PF00138; lectin\_legA; 1.  
DR Pfam; PF00139; lectin\_legB; 1.  
DR ProDom; PD000671; Lectin\_legA; 1.  
DR ProDom; PD000711; Lectin\_legB; 1.  
DR PROSITE; PS00307; LECTIN\_LEGUME\_BETA; 1.  
DR PROSITE; PS00308; LECTIN\_LEGUME\_ALPHA; 1.  
KW Lectin; Calcium; Manganese; Glycoprotein; Acetylation.  
FT CHAIN 1 105  
FT BETA CHAIN.  
FT CHAIN 106 237  
FT CARBOHYD N-LINKED (GLCNAC... ) (POTENTIAL).  
FT MOD\_RES 69 69  
FT MOD\_RES 106 106  
FT METAL 115 115  
FT METAL 117 117  
FT METAL 120 120  
FT METAL 122 122  
FT METAL 127 127  
FT METAL 132 132  
SQ SEQUENCE 237 AA; 25718 MW; BEB7E84DC2895327 CRC64;

Query Match 72.0%; Score 987; DB 1; Length 237;

Best Local Similarity 82.2%; Pred. No. 5.1e-71;

Matches 208; Conservative 3; Mismatches 16; Indels 26; Gaps 7;

Qy 1 AGSLSFSTKFDPNQEDLIFQGHATSTNNVLQTKDSAGNPVSSSAGRVLYSAPRLWE 60

Db 1 AQSLSFSTKFDPNQEDLIFQGTATS-----KLDSDAGNPVSSSAGRVLYSAPRLWE 52

Qy 61 DSAVLTSFDTIINFETISTPTSRITADGLAFFAPPDVSIVSHGGFLGFPNANTLNSST 120

53 DSAVLTSEDPPTI--YIFNTYSRIADGLA-FLAPDPSVISYHGGFLGFPNA----- 101  
 121 SENQTTTAASSNVAVEDT--YLPNDYGDPNYIHGIDVNSIRSKVTAKWQNGKFIAT 179  
 102 ----AESGTAESNVAVEDT--YLPNDYGDPNYIHGIDVNSIRSKVTASWQNGKFIAT 157  
 180 AHISYNSVSKRLSVTSYAG--SKPATLSYDIELHTVLPWVRVGLSASTGQDKERTVHS 238  
 158 AHISYNSVSKRLSVTYYGPKPAT--SYDIELHTVLPWVRVGLSASTGQNIERTVHS 216  
 239 WSFTSSLWTVNAK 251  
 217 WSFTSSLWTVNAK 229

## RESULT 2

LEC1\_C1ALU STANDARD; PRT; 293 AA.  
 ID LEC1\_C1ALU Q39528;

DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Agglutinin I precursor (Clal) (LecClal).  
 OS Cladrastis lutea (yellow wood).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Cladrastis.  
 OX NCBI\_taxID=38412;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 36-55 AND 162-181.  
 RC TISSUE=Bark;  
 RX MEDLINE=96123235; PubMed=8534854;  
 RA van Damme E.J.M., Barre A., Bemer V., Rouge P., van Leuven F.,  
 RA Peumans W.J.;  
 RT "A lectin and a lectin-related protein are the two most prominent  
 proteins in the bark of yellow wood (Cladrastis lutea).";  
 RL Plant Mol. Biol. 29:579-598(1995).  
 CC -!- FUNCTION: MANNOSE/GLUCOSE BINDING BARK LECTIN.  
 CC -!- FUNCTION: BARK LECTINS ARE STORAGE PROTEIN THAT PROBABLY MAINTAINS  
 CC STOCKS OF NITROGEN DURING DORMANT PERIOD. SELF-AGGREGATABLE  
 CC MOLECULES THAT CAN BIND THEIR OWN CARBOHYDRATE SIDE CHAINS. THEY  
 CC COULD ALSO PLAY A ROLE IN THE PLANT'S DEFENCE AGAINST PHYTOPHAGOUS  
 CC INVERTEBRATES OR HERBIVOROUS HIGHER ANIMALS.  
 CC -!- SUBUNIT: HOMOTETRAMER OF FOUR 32 KDA MONOMERS WHICH ARE POST-  
 CC TRANSLATIONALLY CLEAVED INTO A TWO SUBUNITS: A AND B.  
 CC -!- MISCELLANEOUS: BINDS ONE MANGANESE (OR OTHER TRANSITION METAL) ION  
 CC AND ONE CALCIUM ION. THE METAL IONS ARE ESSENTIAL FOR THE  
 CC SACCHARIDE-BINDING AND CELL-AGGLUTININATING ACTIVITIES (BY  
 CC SIMILARITY).  
 CC -!- SIMILARITY: TO OTHER LEGUMINOUS LECTINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U21958; AAC49136.1; -;  
 DR HSP; P19588; 1LUL.  
 DR InterPro; IPR000985; Lectin\_legA.  
 DR InterPro; IPR001220; Lectin\_legB.  
 DR Pfam; PF00138; lectin\_legA; 1.  
 DR Pfam; PF00139; lectin\_legB; 1.  
 DR ProDom; PD000671; Lectin\_legA; 1.  
 DR ProDom; PD000711; Lectin\_legB; 1.  
 DR ProSite; PS00307; LECTIN\_LEGUME\_BETA; 1.  
 DR PROSITE; PS00308; LECTIN\_LEGUME\_ALPHA; 1.  
 KW Lectin; Calcium; Manganese; Glycoprotein; Signal; Mannose-binding.  
 FT SIGNAL 1 35  
 CHAIN 36 161 AGGLUTININ I, SUBUNIT A.

FT CHAIN 162 293 AGGLUTININ I, SUBUNIT B.  
 FT METAL 171 171 MANGANESE (BY SIMILARITY).  
 FT METAL 173 173 MANGANESE AND CALCIUM (BY SIMILARITY).  
 FT METAL 175 175 CALCIUM (BY SIMILARITY).  
 FT METAL 177 177 CALCIUM (BY SIMILARITY).  
 FT METAL 181 181 MANGANESE AND CALCIUM (BY SIMILARITY).  
 FT METAL 186 186 MANGANESE (BY SIMILARITY).  
 FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 47 47 P -> S (IN REF. 1; AA SEQUENCE).  
 FT CONFLICT 162 162 N -> D (IN REF. 1; AA SEQUENCE).  
 SQ SEQUENCE 293 AA; 32128 MW; 48.0% Score 657.5; DB 1; Length 293;

Query Match 48.0%; Score 657.5; DB 1; Length 293;

Best Local Similarity 53.2%; Pred. No. 6.4e-45;

Matches 140; Conservative 34; Mismatches 80; Indels 9; Gaps 6;

QY 3 SLSEFTKFDPNQEDLIFQGHAT--STNNVLTQKLDSSAGNPVSSSAGRVLSAPRLWMD 61  
 DB 38 SLSEFTNNFPNSEDLLIFQKDSISSNETLETRISSGQATSSVGRALYTPVRLWMDK 97  
 QY 62 S-AVLTSFDTIINFELISUPYTSRIADGLAFLAPDPSVISYHGGFLGFPNANLNNSST 120  
 DB 98 STGLASEFKTTFSEAITSP--TQDPDGFAPFIAPDTPPGYGGLLGLF---NGFNLRNS 153  
 QY 121 SENQTTTAASSNVAVEDT--YLPNDYGDPNYIHGIDVNSIRSKVTAKWQNGKFIAT 180  
 DB 154 SNGGAVANNVNSAQVAVEDT--YLPNDYGDPNYIHGIDVNSIRSKVTAKWQNGKFIAT 212  
 QY 181 HISYNSVSKRLSVTSYAG--SKPATLSYDIELHTVLPWVRVGLSASTGQDKERTVHS 240  
 DB 213 QISYNPASQKLTAVTSPNPNPLVSLDIDLTQVLPWVRVGLSASTGQNVNRSILAWS 272  
 QY 241 FTSSLWTVNAKKENKYYITRGV 263  
 DB 273 FSSSLTTLTAKK--EDMYIARVV 293

## RESULT 3

CONA\_CANGL STANDARD; PRT; 290 AA.  
 ID CONA\_CANGL  
 AC P14894;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Concanavalin A precursor (Con A).  
 OS Concanavalin A (Sword bean) (Japanese jack bean).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Canavalia.  
 OX NCBI\_taxID=3824;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Seedling;  
 RX MEDLINE=90127395; PubMed=2404793;  
 RA Yamauchi D., Minamikawa T.;  
 RT "Structure of the gene encoding concanavalin A from Canavalia  
 gladiata and its expression in Escherichia coli cells.";  
 RL FEBS Lett. 260:127-130(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Seed;  
 RA Yamauchi D., Nakamura K., Asahi T., Minamikawa T.;  
 RT "Nucleotide sequence of cDNA for concanavalin A from Canavalia  
 gladiata seeds.";  
 RL Plant Cell Physiol. 30:147-150(1989).  
 CC -!- FUNCTION: D-MANNOSE SPECIFIC LECTIN.  
 CC -!- SUBUNIT: HOMOTETRAMER.  
 CC -!- PM: THE MATURE CHAIN CONSISTS OF RESIDUES 164-281 FOLLOWED BY  
 CC 30-148. TO FORM A NATURE CHAIN THE PRECURSOR UNDERGOES FURTHER  
 CC POST-TRANSLATIONAL MODIFICATION AFTER REMOVAL OF THE SIGNAL  
 CC SEQUENCE; CLEAVAGE AFTER ASN AT POSITIONS 148, 163, AND 281 IS





```

||||| 166 TATAHSYNSASKRLSVSSYPVWVDFVLENNVGPDPVVRVGFSAATGTYTQNNI 225
QY 237 HSWSEFTSSL 245
DB 226 LAWSFRSSL 234

RESULT 5
ID CONA_CANEN STANDARD; PRT; 290 AA.
AC P02866;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Concanavalin A precursor (Con A).
OS Canavalia ensiformis (Jack bean) (Horse bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Canavalia.
OX NCBI_TaxID=3823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85086270; PubMed=3965973;
RA Carrington D.M., Auffret A., Hanke D.E.;
RT "Polypeptide ligation occurs during post-translational modification
of concanavalin A."
RN [2]
RP SEQUENCE OF 30-148.
RX MEDLINE=75095622; PubMed=1112813;
RA Wang J.L., Cunningham B.A., Waxdal M.J., Edelman G.M.;
RT "The covalent and three-dimensional structural of concanavalin A. I.
Amino acid sequence of cyanogen bromide fragments F1 and F2."
RN [3]
RP J. Biol. Chem. 250:1490-1502(1975).
RN [4]
RP SEQUENCE OF 164-281.
RX MEDLINE=75095623; PubMed=1112814;
RA Cunningham B.A., Wang J.L., Waxdal M.J., Edelman G.M.;
RT "The covalent and three-dimensional structure of concanavalin A. II.
Amino acid sequence of cyanogen bromide fragment F3."
RN [5]
RP J. Biol. Chem. 250:1503-1512(1975).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE=73053316; PubMed=4638345;
RA Hardman K.D., Ainsworth C.F.;
RT "Structure of concanavalin A at 2.4-A resolution."
RN [7]
RP Biochemistry 11:4910-4919(1972).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=75095624; PubMed=1112815;
RA Becker J.W., Reeke G.N. Jr., Wang J.L., Cunningham B.A., Edelman G.M.;
RT "The covalent and three-dimensional structure of concanavalin A. III.
Structure of the monomer and its interactions with metals and
saccharides."
RN [9]
RP J. Biol. Chem. 250:1513-1524(1975).
RN [10]
RP X-RAY CRYSTALLOGRAPHY.
RX MEDLINE=75095625; PubMed=1112816;
RA Reeke G.N. Jr., Becker J.W., Edelman G.M.;
RT "The covalent and three-dimensional structure of concanavalin A. IV.
Atomic coordinates, hydrogen bonding, and quaternary structure."
RN [11]
RP J. Biol. Chem. 250:1525-1547(1975).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (1.2 ANGSTROMS).
RA Parkin S., Rupp B., Hope H.;
RT "Atomic resolution structure of concanavalin A at 120 K."
RN [13]
RP Acta Crystallogr. D 52:1161-1168(1996).
RN [14]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
RX MEDLINE=20347885; PubMed=10748006;
RA Bouckaert J., Dewallef Y., Poortmans F., Wyns L., Loris R.;
RT "The structural features of concanavalin A governing non-proline
peptide isomerization."
J. Biol. Chem. 275:19778-19787(2000).
-|- FUNCTION: D-mannose specific lectin.
-|- SUBUNIT: Homotetramer.
-|- PFM: THE MATURE CHAIN CONSISTS OF RESIDUES 164-281 FOLLOWED BY
30-148. TO FORM A MATURE CHAIN THE PRECURSOR UNDERGOES FURTHER
POST-TRANSLATIONAL MODIFICATION AFTER REMOVAL OF THE SIGNAL
SEQUENCE; CLEAVAGE AFTER ASN AT POSITIONS 148, 163, AND 281 IS
FOLLOWED BY TRANSPOSITION AND LIGATION (BY FORMATION OF A NEW
PEPTIDE BOND) OF RESIDUES 164-281 AND 30-148.
-|- MISCELLANEOUS: BINDS ONE MANGANESE (OR OTHER TRANSITION METAL) ION
AND ONE CALCIUM ION. THE METAL IONS ARE ESSENTIAL FOR THE
SACCHARIDE-BINDING AND CELL-AGGLUTININATING ACTIVITIES.
-|- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
-|- DATABASE: NAME=worthington enzyme manual;
WWW="http://www.worthington-biochem.com/manual/C/CONA.html".
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X01632; CAA25787.1; -.
PIR; A03357; CVJBP.
PIR; A03358; CVJJB.
PDB; 1CNI; 30-SEP-83.
PDB; 2CNA; 31-JUL-94.
PDB; 3CNA; 15-APR-91.
PDB; 5CNA; 15-OCT-94.
PDB; 1CON; 31-JAN-94.
PDB; 2CTV; 31-OCT-93.
PDB; 1SCR; 15-OCT-94.
PDB; 1SCR; 15-OCT-94.
PDB; 1APN; 03-APR-96.
PDB; 1CES; 15-FEB-97.
PDB; 1CJP; 15-OCT-97.
PDB; 1CVN; 14-OCT-96.
PDB; 1ENQ; 17-AUG-96.
PDB; 1ENR; 17-AUG-96.
PDB; 2ENR; 16-FEB-99.
PDB; 1ENS; 17-AUG-96.
PDB; 1GIC; 20-AUG-97.
PDB; 1JBC; 12-FEB-97.
PDB; 1NLS; 26-NOV-97.
PDB; 1ONA; 17-SEP-97.
PDB; 1TEI; 24-JUN-98.
PDB; 1VAL; 11-JAN-97.
PDB; 1VAM; 11-JAN-97.
PDB; 1VLN; 01-APR-97.
PDB; 1BXH; 07-OCT-98.
PDB; 2CAU; 25-NOV-98.
PDB; 2CAV; 25-NOV-98.
PDB; 1DQ0; 19-JAN-00.
PDB; 1DQ1; 19-JAN-00.
PDB; 1DQ2; 19-JAN-00.
PDB; 1DQ4; 19-JAN-00.
PDB; 1DQ5; 19-JAN-00.
PDB; 1DQ6; 19-JAN-00.
InterPro; IPR000985; Lectin_lega.
InterPro; IPR001220; Lectin_legB.
Pfam; PF00138; lectin_legA; 1.
Pfam; PF00139; lectin_legB; 1.
ProDom; PD000671; Lectin_legA; 1.
ProDom; PD000711; Lectin_legB; 1.
PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
Lectin; Calcium; Manganese; Glycoprotein; Signal; 3D-structure.
SIGNAL
CHAIN
FT 1 29
FT 30 148 CONCAVALIN (SECOND PART).
FT 149 163
FT PROPEP

```



Db 204 ATAISYNSDFKLSVSSYPNTQANEDYTVSYDVLDTLPEWVRVGFSGTGCVQNH 263  
 QY 235 TVHNSFTSSLTWNVAKKENKXITRGV 263.  
 Db 264 NILSWTFNSLQSSRAKK--EDIYIKRYV 290

RESULT 7  
 LECS\_VATMA STANDARD; PRT; 240 AA.  
 ID LECR\_CLALU  
 AC P81371;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Seed lectin (VML).  
 OS Vatairea macrocarpa.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Dalbergieae; Vatairea.  
 OX NCBI\_TaxID=77050;  
 RN [1]

RP SEQUENCE.  
 RC TISSUE=Seed;  
 RX MEDLINE=98218569; PubMed=9559667;  
 RA Calvete J.J., Santos C.F., Mann K., Grangeiro T.B., Nimtz M.,  
 RA Urbanke C., Sousa-Cavada B.;  
 RT "Amino acid sequence, glycan structure, and proteolytic processing of  
 the lectin of Vatairea macrocarpa seeds.";  
 RL FEBS Lett. 425:286-292(1998).  
 CC -1- FUNCTION: LECTIN THAT BINDS GALACTOSE.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- TISSUE SPECIFICITY: SEED.  
 CC -1- PTM: PARTIALLY N-GLYCOSYLATED AT POSITIONS 111 AND 183 WITH  
 CC THE HEPTASACCHARIDE [(BETA-Xylosyl-1,2)(ALPHA-Mannosyl-1,6)(ALPHA-  
 CC Mannosyl-1,3)]BETA-Mannosyl-1,4-Glcnac-Beta-1,4-Glcnac-Beta-1,4-  
 CC (ALPHA-Fucosyl-1,3)Glcnac. A SMALL PROPORTION OF ALPHA CHAINS ARE  
 CC PROTEOLYTICALLY CLEAVED AT 114-115 INTO GAMMA AND BETA CHAINS.  
 CC THIS IS PROBABLY DEPENDENT ON THE DEGLYCOSYLATION OF N-111.  
 CC -1- MISCELLANEOUS: BINDS ONE MANGANESE (OR OTHER TRANSITION METAL) ION  
 CC AND ONE CALCIUM ION. THE METAL IONS ARE ESSENTIAL FOR THE  
 CC SACCHARIDE-BINDING AND CELL-AGGLUTINATING ACTIVITIES (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.  
 DR HSP; P19588; 110L.

DR GlycoSuiteDB; P81371;  
 DR InterPro: IPR000985; Lectin\_legA.  
 DR InterPro: IPR001220; Lectin\_legB.  
 DR Pfam: PF00138; lectin\_legA; 1.  
 DR Pfam: PF00139; lectin\_legB; 1.  
 DR ProDom: PD000671; Lectin\_legA; 1.  
 DR ProDom: PD000711; Lectin\_legB; 1.  
 DR PROSITE: PS00307; LECTIN\_LEGUME\_BETA; 1.  
 DR PROSITE: PS00308; LECTIN\_LEGUME\_ALPHA; 1.  
 KW Lectin; Calcium; Manganese; Glycoprotein.  
 FT CHAIN 1 240 SEED LECTIN ALPHA CHAIN.  
 FT CHAIN 1 114 SEED LECTIN GAMMA CHAIN.  
 FT CHAIN 115 239 SEED LECTIN BETA CHAIN.  
 FT METAL 123 123 MANGANESE (BY SIMILARITY).  
 FT METAL 125 125 MANGANESE AND CALCIUM (BY SIMILARITY).  
 FT METAL 129 129 CALCIUM (BY SIMILARITY).  
 FT METAL 132 132 MANGANESE AND CALCIUM (BY SIMILARITY).  
 FT METAL 137 137 MANGANESE (BY SIMILARITY).  
 FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .).  
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .).  
 FT VARIANT 117 117 I -> V.  
 FT VARIANT 148 148 M -> K.  
 FT VARIANT 154 154 G -> A.  
 FT VARIANT 168 168 E -> Q.  
 FT UNSURE 239 240  
 SQ SEQUENCE 240 AA; 26197 MW; C17DF6B2568C65C1 CRC64;

Query Match  
 36.7%; Score 502.5; DB 1; Length 240;

Best Local Similarity 44.0%; Pred. No. 8.4e-33;  
 Matches 109; Conservative 43; Mismatches 71; Indels, 25; Gaps 7;  
 QY 4 LSPSFTKFDNQEDLLFOGHATSTN-NVLOVTKLDSAGNPVSSAGRVLYXSAPRLRLWDS 62  
 Db 4 VSFSTKFNPNPNDIIQLQDALVTSKGLQLTKVKD-GKPDVHSLGALYAAPRIHWDS 62  
 QY 63 A-VLTSFDTIINFEISTPYTSRIADGLAFFTAPDPSVISYHGGFLGFLPNANTLNNSSTS 121  
 Db 63 TDRVASFATSFSEFWEAPDESKTADGIAFFLAPDTPQKDDGGFLGFLFNSN----- 114  
 QY 122 ENOTTTKAASSNVVAFEDTYLNPDYGDPNVIHIGIDVNSIRSKVTAKWQNGKIATAH 181  
 Db 115 -----KSQIQTVAVEFDFTFSNT--WDPSARHIGINWNSTESMKYKVGWENGKQVANY 164  
 QY 182 ISYNSVSKRLSVT-SYVAGSKPATLSYDIELHTLVPWVRVGLSASTGQDKERTVH--- 237  
 Db 165 ISYEASTKTLTASLTYPSPNATSIYVSANVDLKSALPEWVRVGFSAATSGLSRHDVETHDVL 224  
 QY 238 SWSFTSSL 245  
 Db 225 DWSFTSVL 232

RESULT 8  
 LECR\_CLALU STANDARD; PRT; 290 AA.  
 ID LECR\_CLALU  
 AC Q39537;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Lectin-related protein precursor (CLERP) (LRPCL) (Fragment).  
 OS Cladrastis lutea (Yellow wood).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Cladrastis.  
 OX NCBI\_TaxID=38412;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 37-56.  
 RC TISSUE=Bark;  
 RX MEDLINE=96123235; PubMed=8534854;  
 RA van Damme E.J.M., Barre A., Bemer V., Rouge P., van Leuven F.,  
 RA Peumans W.J.;  
 RT "A lectin and a lectin-related protein are the two most prominent  
 RT proteins in the bark of yellow wood (Cladrastis lutea).";  
 RL Plant Mol. Biol. 29:579-598(1995).  
 CC -1- FUNCTION: DOES NOT HAVE ANY CARBOHYDRATE BINDING OR AGGLUTINATION  
 CC ACTIVITY.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SIMILARITY: TO OTHER LEGUMINOUS LECTINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U21940; AAC49150.1; -  
 CC HSP; P19588; 110L.  
 DR InterPro: IPR000985; Lectin\_legA.  
 DR InterPro: IPR001220; Lectin\_legB.  
 DR Pfam: PF00138; lectin\_legA; 1.  
 DR Pfam: PF00139; lectin\_legB; 1.  
 DR ProDom: PD000671; Lectin\_legA; 1.  
 DR ProDom: PD000711; Lectin\_legB; 1.  
 DR PROSITE: PS00307; LECTIN\_LEGUME\_BETA; 1.  
 DR PROSITE: PS00308; LECTIN\_LEGUME\_ALPHA; 1.  
 KW Lectin; Calcium; Manganese; Glycoprotein; Signal.  
 FT NON\_TER 1 1  
 FT SIGNAL <1 36  
 FT CHAIN 37 290 LECTIN-RELATED PROTEIN.

DR	Pfam:	PF00139;	lectin_legB; 1.
DR	ProDom:	PD000671;	Lectin_legA; 1.
DR	ProDom:	PD000711;	Lectin_legB; 1.
DR	PROSITE:	PS00307;	LECTIN_LEGUME_BETA; 1.
DR	PROSITE:	PS00308;	LECTIN_LEGUME_ALPHA; 1.
KW	Lectin:	Calcium; Manganese;	Glycoprotein; Signal.
FT	SIGNAL	1	37
FT	CHAIN	38	292
FT	METAL	163	163
FT	METAL	165	165
FT	METAL	167	167
FT	METAL	169	169
FT	METAL	172	172
FT	METAL	177	177
FT	CARBOHYD	82	82
FT	CARBOHYD	154	154
FT	CARBOHYD	186	186
SQ	SEQUENCE	292 AA;	31656 MW; A7431C29117A503E CRC64;
Query Match			
Best Local Similarity 43.3%; Pred. No. 8.7e-32; Length 292;			
Matches 114; Conservative 48; Mismatches 77; Indels 24; Gaps			
QY	4	LSFSFTRDPNQEDLIIPQGHG-TSTNNVLQVKLDSAGNPVSSAGRVLYSAPLRLEWDS 62	
Dd	41	LSFSFPRFASNOEDLLIQGDALLYSKKGELQLTIVEN-GVPINWSTGRALYYAPVHIWDKS 99	
QY	63	-AVLTSEDTTINFESTPYSRTIADGLAFFIAPPDSVI-SYHGFGFLGLFPNANTLNNSST 120	
Dd	100	TGRVASFATSFYVKAPVASKSDGIATFFLAPPNNIOGGHGLGF-----HSS 151	
QY	121	SENOTTTKAASNVAVEFDITLPNDPYDGPNYTHIGIDVNSIRSKVTAKDWONGKIATA 180	
Dd	152	GYN-----SSYQLIAVDFFDHNA--WDENTRHIGIDVNSINSTKTVTWGONGEVANV 203	
QY	181	HISVNSVSKRLSVTSYVAGSKPA-TLSYDIELHTLVPEWVRVGLSASTG---QDKERNV 236	
Dd	204	LISYQAATELTLSLTYPSSQTYSILSAADVLSKILPEWVRVGFTAATGLTTQVVETHDV 263	
QY	237	HSWSFTSSLWTNVAKKENENKYI 259	
Dd	264	LSWSFTSLTGTGCAGADDNVHL 286	
RESULT 10			
LECT_SOPJA			
ID	LECT_SOPJA	STANDARD;	PRT; 270 AA.
AC	P93538;		
DC	15-JUL-1998	(Rel. 36,	Created)
DT	15-JUL-1998	(Rel. 36,	Last sequence update)
DE	30-MAY-2000	(Rel. 39,	Last annotation update)
DE	Bark lectin precursor	(LECSJABG) (Fragment).	
OS	Sophora japonica	(Japanese pagoda tree).	
OC	Eukaryota;	Viridiplantae;	Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta;	Magnoliophyta;	eudicotyledons; core eudicots; Rosidae;
OX	eurosid I;	Fabales;	Fabaceae; Papilionoideae; Sophoreae; Sophora.
NCBI_TaxId=3897;			
[1]			
RN	SEQUENCE FROM N.A.		
RC	TISSUE=Bark;		
RC	MEDLINE=97201486; PubMed=9049272;		
RA	van Damme E.J., Barre A., Rouge P., Peumans W.J.;		
RT	"Molecular cloning of the bark and seed lectins from the Japanese		
RL	pagoda tree ( <i>Sophora japonica</i> ).";		
CC	Plant Mol. Biol. 33:523-536(1997).		
CC	-!- FUNCTION: GALNAC-SPECIFIC LECTIN.		
CC	-!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collabor		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstat		
CC	the European Bioinformatics Institute. There are no restrictions on		
CC	use by non-profit institutions as long as its content is in no		
CC	modified and this statement is not removed. Usage by and for comm		



```

Db 92 STTGNAVSEVTSFSLIKAPNEGKTADGLVFLAPVGTQPLKGGGLGLF----- 142
QY 119 STSENQTTTAASSNVAVVEFTYLNPDYGDPNYTHIGIDVNSIRSKVTAKWDWONGKTA 178
Db 143 ---KDESNK---SQIVAVEEDTFRNVAW-DPENGTHMGIDVNSIQSVTRVRWDWANGEVA 196
QY 179 TAHLSYNSVSKRLSVTSYAG-SKPATLSYDIELHTVLPEWVRVGLSASTGDK---ERN 234
Db 197 NVFTSYEASTKSLTASLVYPSLEKSFILSAIVDLKKVLPEWVRVGTFTAITGLSEDYVQTN 256
QY 235 TVHSWSTSSL--WTNVAKKEN 254
Db 257 DVLWSFESNLPGGNSVASVKN 278

RESULT 12
LEC_SOYBN STANDARD; PRT; 285 AA.
AC P05046;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Lectin precursor (Agglutinin) (SBA).
GN LEL.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=84026469; PubMed=6313203;
RA Vodkin L.O., Rhodes P.R., Goldberg R.B.;
RT "ca lectin gene insertion has the structural features of a
RT transposable element.";
RT Cell 34:1023-1031(1983).
RN [2]
RN X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX MEDLINE=95226380; PubMed=7711015;
RA Dessen A., Gupta D., Sabesan S., Brewer C.F., Sacchettini J.C.;
RT "X-ray crystal structure of the soybean agglutinin cross-linked with
RT a biotennary analog of the blood group I carbohydrate antigen.";
RL Biochemistry 34:4933-4942(1995).
RN [3]
RN X-RAY CRYSTALLOGRAPHY (2.52 ANGSTROMS).
RX MEDLINE=98060759; PubMed=9398234;
RA Olsen L.R., Dessen A., Gupta D., Sabesan S., Sacchettini J.C.,
RA Brewer C.F.;
RT "X-ray crystallographic studies of unique cross-linked lattices
RT between four isomeric biantennary oligosaccharides and soybean
RT agglutinin.";
RL Biochemistry 36:15073-15080(1997).
CC -!- FUNCTION: BINDS GALNAc AND GALACTOSE.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; K00821; AAA33983.1; -.
DR PIR; S27365; S27365.
DR PDB; 2SBA; 09-DEC-98.
DR PDB; 1SBD; 22-APR-98.
DR PDB; 1SBE; 22-APR-98.
DR PDB; 1SBF; 22-APR-98.
DR GlycoSuiteDB; P05046; -.
DR InterPro; IPR000985; Lectin_legA.

```

```

DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
KW Lectin; Glycoprotein; Signal; 3D-structure.
FT SIGNAL 1 32
FT CHAIN 33 285 LECTIN.
FT CARBOHYD 107 107 N-LINKED (GLCNAC...).
SQ SEQUENCE 285 AA; 30928 MW; B3704533C9315C52 CRC64;

Query Match 33.8%; Score 463.5; DB 1; Length 285;
Best Local Similarity 42.8%; Pred. No. 1.2e-29;
Matches 107; Conservative 39; Mismatches 81; Indels 23; Gaps 7;

QY 1 AGSLSFSTKFDPNQEDLIFQGHATSTNN-VLQVTKLDSAGNPVSSSAGRVLYSAPLRW 59
Db 33 AETVSFWMKFPKPNMILQGDIVTSSGKQLNKVDENGTPKPSLGRALYSTPIHIW 92
QY 60 E-DSAVLTISFDITINFEISTPTYSRIADGLAFFIAPPDSVISYHGGFLGFPNANTLNS 118
Db 93 DKETGSVASEFAASFNETFYAPDTKRLADGLAFPLADITKPKQHAGYGLF----- 143
QY 119 STSENQTTTAASSNVAVVEFTYLNPDYGDPNYTHIGIDVNSIRSKVTAKWDWONGKIA 178
Db 144 ---KDESNK---SQIVAVEEDTFRNVAW-DPENGTHMGIDVNSIRSKVTAKWDWONGKIA 194
QY 179 TAHLSYNSVSKRLSVTSYAGSKPATLSYD-IELHTVLPEWVRVGLSASTGDK---KERNT 235
Db 195 KVLIIYDASTLSLVASIVYPSQTSNLSLDVLDKTSLEPWEVRIGFSAATGLDIPGESH 254
QY 236 VHSWSTSSL 245
Db 255 VLSWSFASNL 264

RESULT 13
LEC4_GRISI STANDARD; PRT; 243 AA.
AC P24146;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Lectin IV (GS4).
OS Griffonia simplicifolia.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Caesalpinoideae; Cerciloideae;
OC Griffonia.
OX NCBI_TaxID=3850;
RN [1]
RN SEQUENCE.
RX MEDLINE=93240544; PubMed=8478943;
RA Smillie L.B.;
RL Unpublished results, cited by:
RL Delbaere L.T.J., Vondonselaar M., Prasad L., Quail J.W.,
RL Wilson K.S., Dauter Z.,
RL J. Mol. Biol. 230:950-965(1993).
RN [2]
RN SEQUENCE OF 1-62.
RC TISSUE-Seed;
RX MEDLINE=91097444; PubMed=2268264;
RA Nikrad P.V., Pearlstone J.R., Carpenter M.R., Lemieux R.U.,
RA Smillie L.B.;
RT "Molecular-mass heterogeneity of Griffonia simplicifolia lectin IV
RT subunits. Differences in the oligosaccharide moieties in the
RT N-terminal region.";
RL Biochem. J. 272:343-350(1990).
RN [3]
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=93240544; PubMed=8478943;

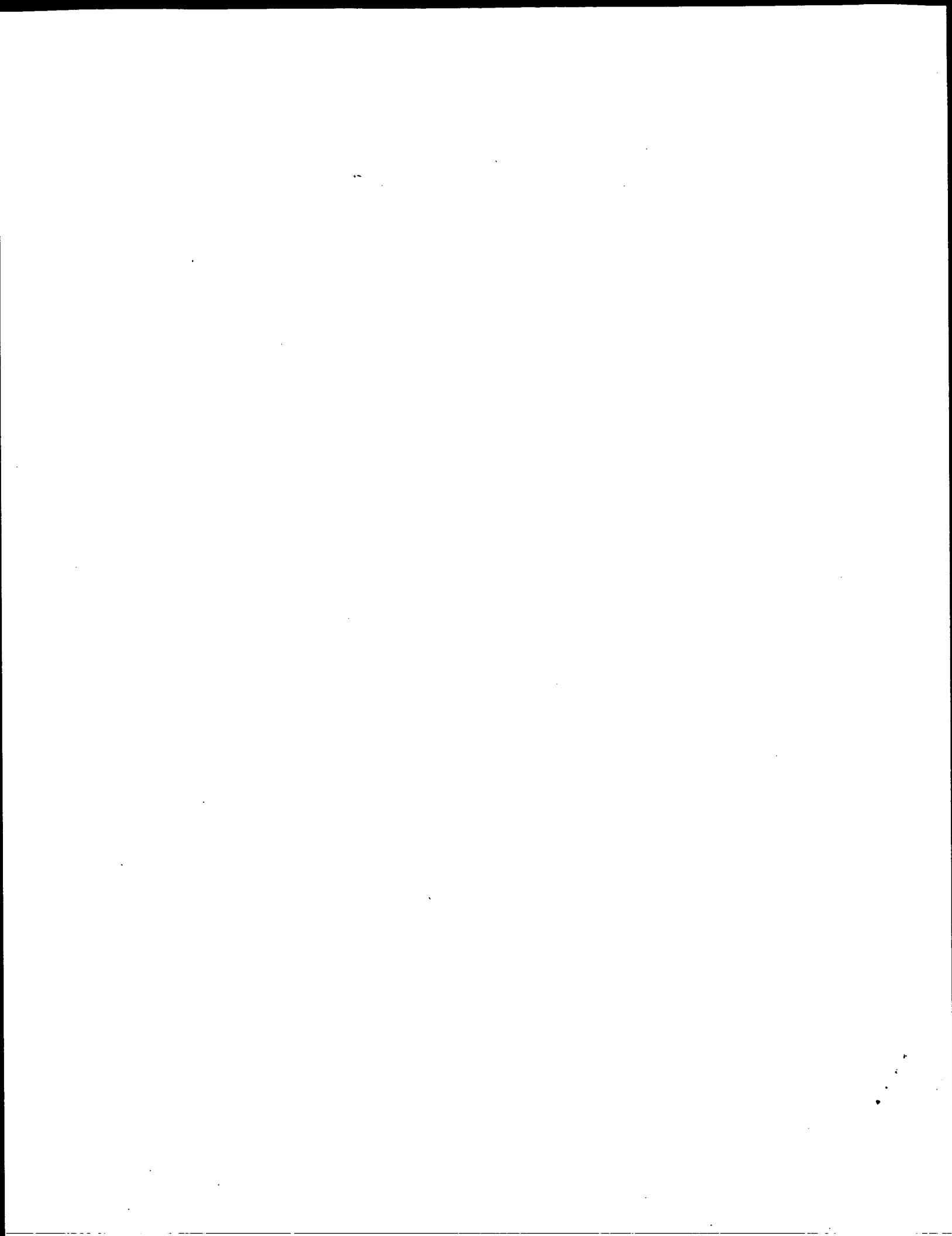
```

Delbaere L.T.J., Vandonselaar M., Prasad L., Quail J.W.,  
 Wilson K.S., Dauter Z.;  
 "Structures of the lectin IV of Griffonia simplicifolia and its  
 complex with the Lewis b human blood group determinant at 2.0-A  
 resolution."; J. Mol. Biol. 230:950-965(1993).  
 CC CC -!- FUNCTION: LECTIN WHICH HAS A STRONG AFFINITY FOR BOTH THE LEWIS B  
 AND Y HUMAN BLOOD-GROUP DETERMINANTS.  
 CC CC -!- SUBUNIT: DIMER OF TWO CHAINS (ALPHA AND BETA) THAT DIFFER ONLY BY  
 THE NUMBER OF OLIGOSACCHARIDE MOTIFETS.  
 CC CC -!- MISCELLANEOUS: BINDS ONE MANGANESE (OR OTHER TRANSITION METAL) ION  
 AND ONE CALCIUM ION. THE METAL IONS ARE ESSENTIAL FOR THE  
 SACCHARIDE-BINDING AND CELL-AGGLUTINATING ACTIVITIES.  
 CC CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.  
 DR DR PIR: S13389; S13389.  
 DR DR PDB: 1LEC; 31-JAN-94.  
 DR DR PDB: 1LED; 31-JAN-94.  
 DR DR PDB: 1GSL; 23-DEC-96.  
 DR DR InterPro: IPR000985; Lectin\_legA.  
 DR DR InterPro: IPR001220; Lectin\_legB.  
 DR DR Pfam: PF00138; Lectin\_legA; 1.  
 DR DR Pfam: PF00139; Lectin\_legB; 1.  
 DR DR ProDom: PD000671; Lectin\_legA; 1.  
 DR DR ProDom: PD000711; Lectin\_legB; 1.  
 DR DR PROSITE: PS00307; LECTIN\_LEGUME\_BETA; 1.  
 DR DR PROSITE: PS00308; LECTIN\_LEGUME\_ALPHA; 1.  
 KW KW Glycoprotein; Lectin; Manganese; Calcium; 3D-structure.  
 FT FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT METAL 129 129 MANGANESE.  
 FT METAL 131 131 MANGANESE AND CALCIUM.  
 FT METAL 133 133 CALCIUM.  
 FT METAL 135 135 CALCIUM.  
 FT METAL 140 140 MANGANESE AND CALCIUM.  
 FT METAL 145 145 MANGANESE.  
 FT CARBOHYD 5 5 N-LINKED (GLCNAC. . .); ONLY IN ALPHA  
 CHAIN.  
 FT CARBOHYD 18 18 N-LINKED (GLCNAC. . .).  
 FT STRAND 5 8  
 FT STRAND 11 11  
 FT TURN 14 15  
 FT TURN 18 20  
 FT STRAND 22 26  
 FT TURN 27 31  
 FT STRAND 29 31  
 FT TURN 32 32  
 FT STRAND 33 36  
 FT STRAND 40 40  
 FT TURN 42 43  
 FT STRAND 46 46  
 FT STRAND 50 56  
 FT STRAND 60 61  
 FT STRAND 63 63  
 FT TURN 65 66  
 FT STRAND 69 69  
 FT STRAND 71 82  
 FT STRAND 88 88  
 FT STRAND 90 97  
 FT TURN 98 99  
 FT HELIX 106 108  
 FT TURN 109 111  
 FT TURN 114 118  
 FT HELIX 120 122  
 FT STRAND 126 131  
 FT HELIX 136 138  
 FT TURN 139 139  
 FT STRAND 145 150  
 FT STRAND 157 160  
 FT HELIX 163 166  
 FT TURN 167 167  
 FT STRAND 171 179  
 FT TURN 180 181  
 FT STRAND 184 191  
 FT TURN 192 193

FT STRAND 196 202  
 FT HELIX 205 208  
 FT STRAND 212 221  
 FT STRAND 223 223  
 FT STRAND 225 237  
 FT TURN 239 242  
 SQ SEQUENCE 243 AA; 26810 MW; CFEC0FA389BBF25 CRC64;  
 Query Match 33.5%; Score 459.5; DB 1; Length 243;  
 Best Local Similarity 44.3%; Pred. No. 2.1e-29;  
 Matches 104; Conservative 33; Mismatches 81; Indels 17; Gaps 4;  
 QY 14 NOEDLIFQGHATSTNNVQLTKDSAGNPVSSAGRVLYSAPRLWEDSAVLTSDTTIN 73  
 DB 18 NGTEITFLGDATRIPLGALQLTKTDANGNPVRSQAQSYSEPVFLWDSTGKAASFTST 77  
 QY 74 FEISTPYTSRIADGLAFIAPDSDVISVHGFGFLGFPNANTLNNSSTSENOTTTKAASN 133  
 DB 78 FLKLN-YCAPTADGLAFIAPVDSVVDYGGFLGF-----RHETADPDKNQ 124  
 QY 134 VVAVFDTYLNPDYGDPNYIHIGIDVNSIRSKVTAKW---DWQNGKIATAHISYNSVKR 190  
 DB 125 VVAVFDTWINKDNDPPYPHIGIDVNSIVSVATTWENDDAYGSSIAHTIYDARSKI 184  
 QY 191 LSVTSYTAGSKPATLSYDIEHLHTVLPWVRVGLSASTGQDKERNVTHWSWSTSL 245  
 DB 185 LTVLLSYEHRDYLILSHVVDLAKVLPQKVRIGFSAGVGYD-EVTYILSMWHFFSTL 238  
 RESULT 14  
 LEC2\_MEDTR STANDARD; PRT; 280 AA.  
 AC Q01807;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Truncated lectin 2 precursor.  
 GN LEC2.  
 OS Medicago truncatula (Barrel medic).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.  
 OX NCBI\_TaxID=3880;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Jemalong;  
 RX MEDLINE=92379255; PubMed=15111126;  
 RA Baichrowitz M.A., Barker D.G., Nadaud I., Rouge P.; Lescure B.;  
 RT "Lectin genes from the legume Medicago truncatula."; Plant Mol. Biol. 19:1011-1017(1992).  
 CC -!- MISCELLANEOUS: LEC2 IS PROBABLY NON FUNCTIONAL, SINCE A FRAMESHIFT  
 MUTATION LEADS TO PREMATURE TRANSLATION TERMINATION AFTER ONLY 98  
 AA. THE SEQUENCE BELOW IGNORES THIS FRAMESHIFT MUTATION.  
 CC -!- MISCELLANEOUS: BINDS ONE MANGANESE (OR OTHER TRANSITION METAL) ION  
 AND ONE CALCIUM ION. THE METAL IONS ARE ESSENTIAL FOR THE  
 SACCHARIDE-BINDING AND CELL-AGGLUTINATING ACTIVITIES.  
 CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X60387; CAA42938.1; -.  
 DR HSSP: P04122; ILOE.  
 DR InterPro: IPR000985; Lectin\_legA.  
 DR InterPro: IPR001220; Lectin\_legB.  
 DR Pfam: PF00138; Lectin\_legA; 1.  
 DR Pfam: PF00139; Lectin\_legB; 1.  
 DR ProDom: PD000671; Lectin\_legA; 1.  
 DR ProDom: PD000671; Lectin\_legB; 1.







GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:36:06 ; Search time 16 seconds  
(without alignments)  
48.067 Million cell updates/sec

Title: US-09-476-485A-24

Perfect score: 36

Sequence: 1 TNNVLQXT 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR\_73.\*

2: piri.\*

3: piri3.\*

4: piri4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	30	83.3	115	2 A71809	hypothetical prote
2	30	83.3	115	2 D64710	hypothetical prote
3	30	83.3	248	2 T26523	hypothetical prote
4	30	83.3	345	2 A97252	sporulation protei
5	30	83.3	544	2 F81059	YibX/YhjW/YijP/Yjd
6	30	83.3	822	2 T41941	glycoprotein B - h
7	30	83.3	960	2 G84652	probable receptor-
8	29	80.6	311	2 G87273	oxidoreductase iro
9	29	80.6	374	2 A46352	ORF1 protein - chl
10	29	80.6	407	2 T04199	hypothetical prote
11	29	80.6	501	2 T43047	retrovirus-related
12	29	80.6	524	2 H84510	probable receptor-
13	29	80.6	677	2 D87018	probable asparagin
14	29	80.6	1184	1 A34795	s kinesin-related pr
15	28	77.8	163	2 B49233	S fibrial adhesin
16	28	77.8	163	2 S15926	sfas protein precu
17	28	77.8	405	2 H71692	hypothetical prote
18	28	77.8	411	2 F97760	hypothetical prote
19	28	77.8	443	2 G64234	hypothetical prote
20	28	77.8	554	2 S67284	hypothetical prote
21	28	77.8	559	2 E89134	hypothetical prote
22	28	77.8	586	2 T24835	protein F25G6.4 [i
23	27	75.0	131	2 I40656	hypothetical prote
24	27	75.0	134	2 F85362	hypothetical prote
25	27	75.0	186	2 T50403	probable succinate
26	27	75.0	208	2 C71697	hypothetical prote
27	27	75.0	208	2 G97767	hypothetical prote
28	27	75.0	254	2 F97263	lactose phosphotra
29	27	75.0	288	2 A90135	SAM-dependent meth

30 27 75.0 311 2 A57223 pheromone receptor  
31 27 75.0 328 2 B65086 hydrogenase (EC 1.  
32 27 75.0 328 2 A85959 hydrogenase-2 smal  
33 27 75.0 328 2 A91114 hydrogenase-2 smal  
34 27 75.0 328 2 AE0885 hydrogenase-2 smal  
35 27 75.0 337 1 WMBEB2 ribonucleoside-dip  
36 27 75.0 337 1 WMBE32 ribonucleoside-dip  
37 27 75.0 340 1 WMBES7 ribonucleoside-dip  
38 27 75.0 351 2 AH1702 ribonucleoside-dip  
39 27 75.0 370 2 S95339 heat shock transcr  
40 27 75.0 380 2 H90261 hypothethical prote  
41 27 75.0 521 2 E64181 probable cytochrom  
42 27 75.0 592 2 F81417 MCP-domain signal  
43 27 75.0 647 2 T28214 probable nucleosid  
44 27 75.0 669 2 E84852 hypothethical prote  
45 27 75.0 697 2 T39512 hypothethical prote

#### ALIGNMENTS

##### RESULT 1

A71809

hypothetical protein jhp1413 - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori

A:Variety: strain J99

C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 28-Jul-2000

C:Accession: A71809

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.

; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F

Nature 397, 176-180, 1999

A>Title: Genomic sequence comparison of two unrelated isolates of the human gastric p

A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: A71809

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-115 <ARN>

A:Cross-references: GB:AE001564; GB:AE001439; NID:g4156032; PIDN:AAD06992.1; PID:g415

A:Experimental source: strain J99

C:Genetics:

A:Gene: jhp1413

C:Superfamily: Helicobacter pylori hypothetical protein HP1524

Query Match 83.3%; Score 30; DB 2; Length 115;  
Best Local Similarity 100.0%; Pred. No. 9.3;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
Db 93 TNNVLQ 98

##### RESULT 2

D64710

hypothetical protein HP1524 - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 28-Jul-2000

C:Accession: D64710

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey,

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,

A>Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: D64710

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-115 <TOM>

A:Cross-references: GB:AE000650; GB:AE000511; NID:g2314700; PIDN:AAD08566.1; PID:g231

C:Superfamily: Helicobacter pylori hypothetical protein HP1524

Query Match 83.3%; Score 30; DB 2; Length 115;

Best Local Similarity 100.0%; Pred. No. 9.3;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
| | | | |  
Db 93 TNNVLQ 98

## RESULT 3

hypothetical protein Y18D10A.11 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T26523

R: Harris, B.

submitted to the EMBL Data Library, December 1998

A:Reference number: Z20226

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-248 <WIL>

A:Cross-references: EMBL:AL034393; PIDN:CAA22314.1; CESP:Y18D10A.11

A:Experimental source: clone Y18D10A

C:Genetics:

A:Gene: CESP:Y18D10A.11

A:Introns: 55/2; 101/2; 142/2; 204/3

Query Match 83.3%; Score 30; DB 2; Length 248;

Best Local Similarity 85.7%; Pred. No. 22;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNVLOXT 8  
| | | | |  
Db 150 NNVLOXT 156

## RESULT 4

sporulation protein SpoIID [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 30-Sep-2001

C:Accession: A97252

R: Nollong, J.; Braton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-345 <KOR>

A:Cross-references: GB:AE001437; PIDN:AAK80804.1; PID:gl5025907; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC2861

C:Superfamily: stage II sporulation protein D

## RESULT 5

YhbX/YhbW/YajP/YjdB family protein NMB1638 [imported] - Neisseria meningitidis (strain M

C:Species: Neisseria meningitidis

C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Feb-2001

C:Accession: F81059; E81816

R: Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A

QY 2 NNVLOXT 8  
| | | | |  
Db 271 NNVLOXT 277

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.  
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.  
Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755; PMID:10710307

A:Accession: F81059

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-544 <VET>

A:Cross-references: GB:AE002514; GB:AE002098; NID:g7226886; PIDN:AAF41987.1; PID:g722

A:Experimental source: serogroup B, strain MC58

R: Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491

A:Reference number: A81775; MUID:20222556; PMID:10761919

A:Accession: E81816

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-544 <PAR>

A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85114.1; PID:g738

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMB1638; NMB1892

C:Superfamily: Escherichia coli yidB protein

Query Match 83.3%; Score 30; DB 2; Length 544;

Best Local Similarity 85.7%; Pred. No. 54;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNVLOXT 8  
| | | | |  
Db 105 NNVLOXT 111

## RESULT 6

T41941

glycoprotein B - human herpesvirus 7 (strain JI)

C:Species: human herpesvirus 7

A:Variety: strain JI

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000

C:Accession: T41941

R: Nicholas, J.

submitted to the EMBL Data Library, December 1995

A:Description: Determination and analysis of the complete nucleotide sequence of huma

A:Reference number: Z22022

A:Accession: T41941

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-822 <NIC>

A:Cross-references: EMBL:U43400; PIDN:AAC54701.1

A:Experimental source: strain JI

C:Genetics:

A:Note: U39

C:Superfamily: herpesvirus glycoprotein B

## RESULT 7

G84652

probable receptor-like protein kinase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: G84652

R: Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y

Query Match 83.3%; Score 30; DB 2; Length 822;

Best Local Similarity 75.0%; Pred. No. 87;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8  
| | | | |  
Db 726 TNSVLQAT 733

M.; Koo, M.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: G84652

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-960 <STO>

A:Cross-references: GB:AE002093; NID:g3643604; PIDN:AAC42251.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g25790

A:Map position: 2

Query Match 83.3%; Score 30; DB 2; Length 960;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNVLQ 6  
|||||

Db 932 TNNVLQ 937

RESULT 8

G87273

oxidoreductase iron/ascorbate family CC0200 [imported] - *Caulobacter crescentus*

C:Species: *Caulobacter crescentus*

C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C:Accession: G87273

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J. B.; Laub, M.T.; DeBoy, R.F.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of *Caulobacter crescentus*.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: G87273

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-311 <STO>

A:Cross-references: GB:AE005673; NID:g13421323; PIDN:AAK22187.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC0200

Query Match 80.6%; Score 29; DB 2; Length 311;

Best Local Similarity 75.0%; Pred. No. 49;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8  
|||||

Db 238 TNNVLPST 245

RESULT 9

A46352

ORF1 protein - *Chlorella virus PBCV-1*

C:Species: *Chlorella virus PBCV-1*

C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 17-Nov-2000

C:Accession: A46352; T18035

R:Schuster, A.M.; Graves, M.; Korth, K.; Ziegelbein, M.; Brumbaugh, J.; Grone, D.; Meint Virology 176, 515-523, 1990

A:Title: Transcription and sequence studies of a 4.3-kbp fragment from a ds-DNA eukaryot

A:Reference number: A46352; MUID:90266467; PMID:2345963

A:Accession: A46352

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-374 <SCH>

A:Cross-references: GB:M33758; NID:g323370; PIDN:AAA66400.1; PID:g807594

R:Graves, M.V.; Van Etten, J.L.

submitted to the EMBL Data Library, May 1999

A:Reference number: Z18806

A:Accession: T18035

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-82, 'T', 84-374 <GRA>

A:Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96900.1

A:Experimental source: specific host *Chlorella* strain NC64A

C:Genetics:

A:Note: A533R

Query Match 80.6%; Score 29; DB 2; Length 374;

Best Local Similarity 85.7%; Pred. No. 61;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NNVLOXT 8  
|||||

Db 112 NNVLOQT 118

RESULT 10

T04199

hypothetical protein T4F9.100 - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 30-Apr-1999

C:Accession: T04199

R:Bevan, M.; Wedler, H.; Wedler, E.; Wambutt, R.; Bancroft, I.; Meves, H.W.; Mayer, K submitted to the Protein Sequence Database, March 1999

A:Reference number: Z15260

A:Accession: T04199

A:Molecule type: DNA

A:Residues: 1-407 <BEV>

A:Cross-references: EMBL:AL049523

A:Experimental source: cultivar Columbia; BAC clone T4F9

C:Genetics:

A:Map position: 4

A:Introns: 37/3; 114/3; 210/3; 251/3; 327/3

A:Note: T4F9.100

Query Match 80.6%; Score 29; DB 2; Length 407;

Best Local Similarity 75.0%; Pred. No. 67;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8  
|||||

Db 54 TNSVLQHT 61

RESULT 11

T43047

retrovirus-related env polyprotein homolog - Mediterranean fruit fly retrotransposon

C:Species: *Ceratitis capitata* (Mediterranean fruit fly)

C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 11-Jan-2000

C:Accession: T43047

R:Zhou, Q.; Haymer, D.S.

submitted to the EMBL Data Library, June 1996

A:Description: Gypsy-like retrotransposon in the Medfly.

A:Reference number: Z22299

A:Accession: T43047

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-501 <ZHO>

A:Cross-references: EMBL:U60529; NID:g1402846; PID:g1402849; PIDN:AAC28744.1

A:Experimental source: strain Med+; clone B

C:Genetics:

A:Mobile element: retrotransposon yoyo

Query Match 80.6%; Score 29; DB 2; Length 501;

Best Local Similarity 83.3%; Pred. No. 84;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNVLQ 6  
|||||

Db 147 TNNILQ 152

RESULT 12

H84510

probable receptor-like protein kinase [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: H84510  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: H84510  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-524 <STO>  
 A:Cross-references: GB:AE002093; NID:g4726119; PIDN:AAD28319.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g13800  
 A:Map position: 2

Query Match 80.6%; Score 29; DB 2; Length 524;  
 Best Local Similarity 83.3%; Pred. No. 89;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TNNVLQ 6  
 |||||  
 Db 44 TNNILQ 49

RESULT 13  
 D87018  
 C:Species: Mycobacterium leprae  
 Probable asparagine synthetase [imported] - Mycobacterium leprae  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
 C:Accession: D87018  
 R:Cole, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho  
 R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,  
 eam, M.A.; Rutherford, K.M.  
 Nature 409, 1007-1011, 2001  
 A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq  
 A:Title: Massive gene decay in the leprosy bacillus.  
 A:Reference number: A86909; MUID:21128732; PMID:11234002  
 A:Accession: D87018  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-677 <STO>  
 A:Cross-references: GB:AL450380; NID:g13092947; PIDN:CAC31255.1; GSPDB:GN00147  
 C:Genetics:  
 A:Gene: asnB

Query Match 80.6%; Score 29; DB 2; Length 677;  
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
 |||||  
 Db 39 TNNILQ 44

RESULT 14  
 A34795  
 Kinesin-related protein bimC - Emericella nidulans  
 C:Species: Emericella nidulans, Aspergillus nidulans  
 C:Date: 03-Mar-1994 #sequence\_revision 03-Mar-1994 #text\_change 19-Jan-2001  
 C:Accession: A34795; B34795  
 R:Enos, A.P.; Morris, N.R.  
 Cell 60, 1019-1027, 1990  
 A:Title: Mutation of a gene that encodes a kinesin-like protein blocks nuclear division  
 A:Reference number: A34795; MUID:90199865; PMID:2138511  
 A:Accession: A34795  
 A:Molecule type: DNA  
 A:Residues: 1-1184 <ENO>  
 A:Cross-references: GB:M32075; NID:g168022; PIDN:AAA33298.1; PID:g168023  
 C:Superfamily: Kinesin-related protein Eg5; kinesin motor domain homology

C:Keywords: ATP; microtubule binding; mitosis; nucleotide binding; P-loop  
 F:82-422/Domain: kinesin motor domain homology <KNOT>  
 F:167-174/Region: nucleotide-binding motif A (P-loop)  
 F:173/Binding site: ATP (Lys) #status predicted  
 Query Match 80.6%; Score 29; DB 1; Length 1184;  
 Best Local Similarity 75.0%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 TNNVLQXT 8  
 |||||  
 Db 521 TNDVLQQT 528  
 RESULT 15  
 B49233  
 S fimbrial adhesin minor subunit sfas - Escherichia coli  
 C:Species: Escherichia coli  
 C:Date: 19-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 24-Sep-1999  
 C:Accession: B49233  
 R:Hacker, J.; Kestler, H.; Hoshutzky, H.; Jann, K.; Lottspeich, F.; Korhonen, T.K.  
 Infect. Immun. 61, 544-550, 1993  
 A:Title: Cloning and characterization of the S fimbrial adhesin II complex of an Esch  
 A:Reference number: A49233; MUID:93138776; PMID:8093693  
 A:Contents: O18:K1  
 A:Accession: B49233  
 A:Status: preliminary  
 A:Molecule type: nucleic acid  
 A:Residues: 1-163 <HAC>  
 A:Cross-references: GB:S53210; NID:g264033; PIDN:AB25046.1; PID:g264035  
 A:Note: sequence extracted from NCBI backbone (NCBIN:123181, NCBIP:123183)  
 C:Superfamily: conserved hypothetical protein bl503

Query Match 77.8%; Score 28; DB 2; Length 163;  
 Best Local Similarity 75.0%; Pred. No. 40;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8  
 |||||  
 Db 30 TGNVLQRT 37

Search completed: February 26, 2003, 15:38:35  
 Job time : 18 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:33:05 ; Search time 29 Seconds  
(without alignments)  
11.442 Million cell updates/sec

Title: US-09-476-485A-24  
Perfect score: 36  
Sequence: 1 TNNVLQXT 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	30	83.3	822	1 VGLB_HSV7J	P52352 human herpe
2	30	83.3	1762	1 DPOQ_HUMAN	O75417 homo sapien
3	29	80.6	1184	1 BINC_EMENI	P17120 emericella
4	28	77.8	163	1 SFAS_ECOLI	P13430 escherichia
5	28	77.8	443	1 Y314_MFCGE	Q49415 mycoplasma
6	27	75.0	208	1 Y396_RICPR	Q92dd5 rickettsia
7	27	75.0	328	1 HYBA_ECOLI	P03179 escherichia
8	27	75.0	337	1 RIR2_HSV23	P03174 herpes simp
9	27	75.0	340	1 RIR2_HSV11	P10224 herpes simp
10	27	75.0	340	1 RIR2_HSV1K	P06474 herpes simp
11	27	75.0	365	1 SOXB_RHOSO	P54997 rhodococcus
12	27	75.0	521	1 CYOA_HAEIN	P45021 haemophilus
13	27	75.0	638	1 60IM_COXBU	P45650 coxiella bu
14	27	75.0	647	1 COAT_ADVG	P24029 aleutian m
15	27	75.0	647	1 NTP1_MSEPV	Q9Yw39 melanoplus
16	27	75.0	718	1 GUAA_DICDI	P32073 dictyosteli
17	27	75.0	955	1 SVV_BUCAI	P57447 buchnera ap
18	27	75.0	993	1 DPOL_BPAPS	P971q3 bacterioph
19	27	75.0	2569	1 LMA3_MOUSE	O61789 mus musculu
20	26	72.2	170	1 Y051_UREPA	Q9pr94 ureaplasma
21	26	72.2	274	1 TRV5_ANOGA	P35039 anopheles g
22	26	72.2	283	1 YG20_YEAST	P33068 saccharomyc
23	26	72.2	289	1 MW1_YEAST	P39731 thauera aro
24	26	72.2	324	1 HCRB_THAAR	O33820 homo sapien
25	26	72.2	357	1 O2B2_HUMAN	Q9g2k3 schizosacch
26	26	72.2	423	1 YE18_SCHPO	O13873 schizosacch
27	26	72.2	500	1 LCVB_TOBAC	Q43578 nicotiana t
28	26	72.2	522	1 MAL2_DROME	P07190 drosophila
29	26	72.2	680	1 OPDA_ECOLI	P27298 escherichia
30	26	72.2	680	1 OPDA_SALTY	P27237 salmonella
31	26	72.2	681	1 OPDA_HAEIN	P44573 haemophilus
32	26	72.2	734	1 UN36_CAEEL	P34374 caenorhabdi
33	26	72.2	750	1 YD33_MYCPN	P75445 mycoplasma

34	26	72.2	804	1 VP5_WTV	P12366 wound tumor
35	26	72.2	928	1 PM10_CHLPN	Q9rb65 chlamydia p
36	26	72.2	1064	1 YY08_METJA	Q60307 methanococc
37	26	72.2	1178	1 PHVB_SORBI	P93527 sorghum bic
38	26	72.2	1205	1 YLJ9_CAEEL	P34372 caenorhabdi
39	26	72.2	1233	1 YF16_YEAST	P43597 caenorhabdi
40	26	72.2	1398	1 DNA2_SCHPO	Q9ur02 schizosacch
41	25	69.4	68	1 Y244_RICPR	Q92dt1 rickettsia
42	25	69.4	69	1 GBG7_RAT	P43425 rattus norv
43	25	69.4	109	1 RR10_CVACA	Q9tlv9 cyanidium c
44	25	69.4	113	1 DYX_MOUSE	P51807 mus musculu
45	25	69.4	148	1 Y222_TREPA	O83251 treponema p

ALIGNMENTS

RESULT 1  
VGLB\_HSV7J  
ID VGLB\_HSV7J STANDARD; PRT; 822 AA.  
AC P52352;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Glycoprotein B precursor.  
GN GB OR U39.  
OS Human herpesvirus (type 7 / strain J1) (HHV7).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Simplexvirus.  
OX NCBI\_TaxID=57278;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Nicholas J.;  
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL: U43400; AAC54701.1; -  
CC InterPro: IPR000234; Glycoprot\_B.  
CC Pfam: PF00606; Glycoprotein\_B; 1.  
CC ProDom: PD000693; Glycoprot\_B; 1.  
CC Glycoprotein; Transmembrane; Signal.  
FT SIGNAL 1 22 POTENTIAL.  
FT CHAIN 23 822 GLYCOPROTEIN B.  
FT DOMAIN 23 649 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 650 670 POTENTIAL.  
FT DOMAIN 671 683 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 684 704 POTENTIAL.  
FT DOMAIN 705 822 EXTRACELLULAR (POTENTIAL).  
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 244 244 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 338 338 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 358 358 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 517 517 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 822 AA; 93147 MW; F2B410BD359C6498 CRC64;

Query Match 83.3%; Score 30; DB 1; Length 822;  
Best Local Similarity 75.0%; Pred. No. 24;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 TNNVLOXT 8
Db 726 TNSVLQAT 733

RESULT 2
DPOQ_HUMAN STANDARD; PRT; 1762 AA.
AC O75417; O95160;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase theta (EC 2.7.7.7) (DNA polymerase eta).
GN POLQ OR POLH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA MEDLINE=89326527; PubMed=10395804;
RA Sharief F.S., Vojta P.J., Ropp P.A., Copeland W.C.;
RT "Cloning and chromosomal mapping of the human DNA polymerase theta
RT (POLQ), the eighth human DNA polymerase.";
RL Genomics 59:90-96(1999).
RN [2]
RP SEQUENCE OF 607-1762 FROM N.A.
RA Harris P.V., Kaelin C.B., Burtis K.C.;
RT "Catalytic activity of Pol eta, a new human DNA polymerase related to
RT the bacterial DNA polymerase I family and Drosophila Mus308.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COULD BE INVOLVED IN THE REPAIR OF INTERSTRAND
CC CROSSLINKS.
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA}(N).
CC -!- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF052573; AAC33565.1; -.
CC DR EMBL; AF043628; AAD05272.1; -.
CC DR HSSP; P00582; 1KFS.
CC DR Genew; HGNC:9186; POLQ.
CC DR MIW; 604419; -.
CC DR InterPro; IPR001098; DNA_pol.
CC DR InterPro; IPR002298; DNA_polI.
CC DR Pfam; PF00476; DNA_pol_A; 1.
CC DR PRINTS; PR00868; DNAPOLI.
CC DR SMART; SM00482; POLAC; 1.
CC DR PROSITE; PS00447; DNA_POLYMERASE_A; 1.
KW Transferase; DNA-directed DNA polymerase; ATP-binding; DNA repair.
FT NP_BIND 175 182 ATP (POTENTIAL).
FT CONFLICT 1185 1185 L -> F (IN REF. 2).
FT CONFLICT 1685 1685 R -> Q (IN REF. 2).
FT CONFLICT 1719 1719 V -> A (IN REF. 2).
FT CONFLICT 1756 1762 ELKDFDV -> RAKGL (IN REF. 2).
FT CONFLICT 1762 1762 ELKDFDV -> RAKGL (IN REF. 2).
SQ SEQUENCE 1762 AA; 197596 MW; AEC17A2103F6BDA CRC64;

Query Match 83.3%; Score 30; DB 1; Length 1762;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLO 6
Db 934 TNNVLO 939

RESULT 3
BIMC_EMENI STANDARD; PRT; 1184 AA.
ID BIMC_EMENI
AC P17120;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Kinesin-like protein BIMC.
GN BIMC.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=90199865; PubMed=2138511;
RA Enos A.P., Morris N.R.;
RT "Mutation of a gene that encodes a kinesin-like protein blocks
RT nuclear division in A. nidulans.";
RL Cell 60:1019-1027(1990).
CC -!- FUNCTION: IMPORTANT ROLE IN MITOTIC DIVIDING CELLS. MICROTUBULE
CC MOTOR REQUIRED FOR SPINDLE BODY SEPARATION.
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BIMC
CC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M32075; AAA33298.1; -.
CC DR PIR; A34795; A34795.
CC DR HSSP; P17119; 3KAR.
CC DR InterPro; IPR001752; kinesin_motor.
CC DR Pfam; PF00225; kinesin; 1.
CC DR PRINTS; PR00380; KINESINHEAVY.
CC DR SMART; SM00129; Kisc; 1.
CC DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
CC DR PROSITE; PS00067; KINESIN_MOTOR_DOMAIN2; 1.
KW Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;
KW Mitosis; Cell cycle; Phosphorylation.
FT DOMAIN 79 485 KINESIN-MOTOR (BY SIMILARITY).
FT DOMAIN 489 900 COILED COIL (POTENTIAL).
FT DOMAIN 901 1184 GLOBULAR (POTENTIAL).
FT NP_BIND 167 174 ATP (BY SIMILARITY).
FT MOD_RES 1006 1006 PHOSPHORYLATION (BY CDC2) (BY
FT SIMILARITY).
SQ SEQUENCE 1184 AA; 131630 MW; CEA015EC8F980E4F CRC64;

Query Match 80.6%; Score 29; DB 1; Length 1184;
Best Local Similarity 75.0%; Pred. No. 63;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
Db 521 TNDVLOQT 528

RESULT 4
SFAS_ECOLI STANDARD; PRT; 163 AA.
ID SFAS_ECOLI
AC P13430;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE S-fimbrial adhesin protein sfas precursor.
GN SFAS.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

```



OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RX STRAIN=06:K15;  
 RA MEDLINE=90138121; PubMed=2576095;  
 RA Jann K., Hacker J.;  
 RT "Analysis of genes coding for the stalic acid-binding adhesin and two  
 RT other minor fimbrial subunits of the S-fimbrial adhesin determinant  
 RT of Escherichia coli";  
 RL Mol. Microbiol. 3:1735-1744(1989).  
 RN [2]  
 RP MUTAGENESIS OF LYS-138; ARG-140 AND LYS-144.  
 RX MEDLINE=90307213; PubMed=2194961;  
 RA Morschhaeuser J., Hoschuetzky H., Jann K., Hacker J.;  
 RT "Functional analysis of the stalic acid-binding adhesin sfas of  
 RT pathogenic Escherichia coli by site-specific mutagenesis.";  
 RL Infect. Immun. 58:2133-2138(1990).  
 CC -!- FUNCTION: FIMBRIAE (ALSO CALLED PILI), POLAR FILAMENTS RADIATING  
 CC FROM THE SURFACE OF THE BACTERIUM TO A LENGTH OF 0.5-1.5  
 CC MICROMETERS AND NUMBERING 100-300 PER CELL, ENABLE BACTERIA TO  
 CC COLONIZE THE EPITHELIUM OF SPECIFIC HOST ORGANS.  
 CC -!- FUNCTION: THIS PROTEIN IS THE STALIC ACID-BINDING SUBUNIT OF THE  
 CC S-FIMBRIAL ADHESINS. IT ALSO PLAYS A ROLE IN DETERMINATION OF THE  
 CC DEGREE OF FIMBRATION OF THE CELL.  
 CC -!- DISEASE: S-FIMBRIAL ADHESINS ENABLE PATHOGENIC E.COLI CAUSING  
 CC URINARY-TRACT INFECTIONS OR NEWBORN MENINGITIS TO ATTACH TO  
 CC GLYCOPROTEINS TERMINATING WITH ALPHA-SIALIC ACID-(2-3)-BETA-GAL.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X16664; CAA34653.1; .  
 DR PIR: S06194; S06194.  
 DR PIR: S15926; S15926.  
 DR InterPro: IPR000259; Fimbrial.  
 DR Pfam: PF00419; Fimbrial; 1.  
 KW Fimbria; Signal.  
 FT SIGNAL 1 22  
 FT CHAIN 23 163 S-FIMBRIAL ADHESIN PROTEIN SFAS.  
 FT DISULFID 38 75 PROBABLE.  
 FT DOMAIN 138 144 INVOLVED IN SIALIC ACID BINDING.  
 FT MUTAGEN 138 138 K->T: NO CHANGE IN S-BINDING.  
 FT MUTAGEN 140 140 R->S: NO HEMAGGLUTINATION WEAK REACTION  
 FT WITH ANTIADHESIN-SPECIFIC ANTIBODY A1.  
 FT MUTAGEN 144 144 K->T: NO HEMAGGLUTINATION, NO REACTION  
 FT WITH ANTIADHESIN-SPECIFIC ANTIBODY A1.  
 SQ SEQUENCE 163 AA; 17183 MW; 0BF1333BF8B2DE4F CRC64;  
 Query Match 77.8%; Score 28; DB 1; Length 163;  
 Best Local Similarity 75.0%; Pred. No. 10;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 TNNVLQXT 8  
 Db 30 TGNVLQRT 37  
 RESULT 5  
 Y314\_MYCGE  
 ID Y314\_MYCGE STANDARD; PRT; 443 AA.  
 AC Q49415; Q49279;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein MG314.  
 GN MG314.

OS Mycoplasma genitalium.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2097;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RX STRAIN=ATCC 33530 / G-37;  
 RX MEDLINE=96026346; PubMed=7569993;  
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
 RA Friedman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,  
 RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
 RA Tomb J.-F., Dougherty B.A., Boff K.F., Hu P.-C., Lucier T.S.,  
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
 RT "The minimal gene complement of Mycoplasma genitalium";  
 RL Science 270:397-403(1995).  
 RN [2]  
 RP SEQUENCE OF 70-171 FROM N.A.  
 RC STRAIN=ATCC 33530 / G-37;  
 RX MEDLINE=94075230; PubMed=8253680;  
 RA Peterson S.N., Hu P.-C., Boff K.F., Hutchison C.A. III;  
 RT "A survey of the Mycoplasma genitalium genome by using random  
 RT sequencing.";  
 RL J. Bacteriol. 175:7918-7930(1993).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U39712; AAC71536.1; .  
 DR EMBL: U02151; AAD12432.1; .  
 DR TIGR: MG314; .  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 443 AA; 51132 MW; 3EFB03F9168015D7 CRC64;  
 Query Match 77.8%; Score 28; DB 1; Length 443;  
 Best Local Similarity 71.4%; Pred. No. 34;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 NNVLQXT 8  
 Db 158 NNILQQT 164  
 RESULT 6  
 Y396\_RICPR  
 ID Y396\_RICPR STANDARD; PRT; 208 AA.  
 AC Q92DD5;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein RP396.  
 GN RP396.  
 OS Rickettsia prowazekii.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;  
 OC Rickettsiaceae; Rickettsia;  
 OX NCBI\_TaxID=782;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Madrid E;  
 RX MEDLINE=99039499; PubMed=9823893;  
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,  
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;  
 RT "The genome sequence of Rickettsia prowazekii and the origin of  
 RT mitochondria";  
 RL Nature 396:133-140(1998).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: AJ235271; CAA14853.1; -  
 DR InterPro: IPR000627; Dioxigenase.  
 DR Pfam: PF00775; Dioxigenase; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 208 AA; 23985 MW; 57BB82FADA62F864 CRC64;

Query Match 75.0%; Score 27; DB 1; Length 208;  
 Best Local Similarity 62.5%; Pred. No. 24;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8  
 I I I I I  
 Db 48 TNNLLRKT 55

RESULT 7  
 HYBA\_ECOLI STANDARD; PRT; 328 AA.  
 AC P37179;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hydrogenase-2 operon protein hyba precursor.  
 GN HYBA OR B2996 OR Z4350 OR ECS3881.  
 OS Escherichia coli, and  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562, 83334;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / TGI;  
 RX MEDLINE=94292472; PubMed=8021226;  
 RA Menon N.K., Chatelus C.Y., Dervartanian M., Wendt J.C.,  
 RA Shanmugam K.T., Peck H.D. Jr., Przybyla A.E.;  
 RT "Cloning, sequencing, and mutational analysis of the hyb operon  
 RT encoding Escherichia coli hydrogenase 2";  
 RL J. Bacteriol. 176:4416-4423(1994).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12";  
 RL Science 277:1453-1474(1997).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
 RL Nature 409:529-533(2001).  
 [4]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RIMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Toke T.,  
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,

RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
 RT O157:H7 and genomic comparison with a laboratory strain K-12";  
 RL DNA Res. 8:11-22(2001).  
 CC -!- FUNCTION: PARTICIPATES IN THE PERIPLASMIC ELECTRON-TRANSFERRING  
 CC ACTIVITY OF HYDROGENASE 2 DURING ITS CATALYTIC TURNOVER.  
 CC -!- COFACTOR: BINDS 3 4FE-4S CLUSTERS AND A 3FE-4S CLUSTER.  
 CC -!- SUBCELLULAR LOCATION: Periplasmic.  
 CC -!- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF  
 CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.  
 CC -!- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE THE SMALL SUBUNIT  
 CC OF HYDROGENASE 2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: U09177; AAA21589.1; -  
 DR EMBL: U28377; AAA69163.1; -  
 DR EMBL: AE000382; AAC76032.1; -  
 DR EMBL: AE005529; AAG58133.1; -  
 DR EMBL: AF002563; BAB37304.1; -  
 DR EcoGene; EGI1799; hybA.  
 DR InterPro: IPR001450; 4Fe4S\_ferredoxin.  
 DR Pfam: PF00037; fer4; 1.  
 DR PROSITE: PS00198; 4FE4S\_FERREDOXIN; 1.  
 KW Oxidoreductase; Signal; Periplasmic; Iron-sulfur; 4Fe-4S; 3Fe-4S;  
 KW Complete proteome.  
 FT SIGNAL 1 27 POTENTIAL.  
 FT CHAIN 28 328 HYDROGENASE-2 OPERON PROTEIN HYBA.  
 FT METAL 47 47 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).  
 FT METAL 50 50 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).  
 FT METAL 53 53 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).  
 FT METAL 57 57 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).  
 FT METAL 112 112 IRON-SULFUR 2 (3FE-4S) (POTENTIAL).  
 FT METAL 115 115 IRON-SULFUR 2 (3FE-4S) (POTENTIAL).  
 FT METAL 120 120 IRON-SULFUR 2 (3FE-4S) (POTENTIAL).  
 FT METAL 124 124 IRON-SULFUR 2 (3FE-4S) (POTENTIAL).  
 FT METAL 145 145 IRON-SULFUR 3 (4FE-4S) (POTENTIAL).  
 FT METAL 148 148 IRON-SULFUR 3 (4FE-4S) (POTENTIAL).  
 FT METAL 151 151 IRON-SULFUR 3 (4FE-4S) (POTENTIAL).  
 FT METAL 155 155 IRON-SULFUR 4 (4FE-4S) (POTENTIAL).  
 FT METAL 174 174 IRON-SULFUR 4 (4FE-4S) (POTENTIAL).  
 FT METAL 177 177 IRON-SULFUR 4 (4FE-4S) (POTENTIAL).  
 FT METAL 193 193 IRON-SULFUR 4 (4FE-4S) (POTENTIAL).  
 FT METAL 197 197 IRON-SULFUR 4 (4FE-4S) (POTENTIAL).  
 SQ SEQUENCE 328 AA; 36003 MW; 77203A0F50F61662 CRC64;

Query Match 75.0%; Score 27; DB 1; Length 328;  
 Best Local Similarity 66.7%; Pred. No. 42;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
 I I I I I  
 Db 83 TNNIIQ 88

RESULT 8  
 RIR2\_HSV23 STANDARD; PRT; 337 AA.  
 ID RIR2\_HSV23  
 AC P03174;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)  
 DE (Ribonucleotide reductase) (38 kDa subunit).  
 OS Herpes simplex virus (type 2 / strain 333).  
 OS Herpes simplex virus (type 2 / strain HG52).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

```

OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10313, 10315;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=333;
RX MEDLINE=84138764; PubMed=6321759;
RA Galloway D.A., Swain M.A.;
RT "Organization of the left-hand end of the herpes simplex virus type 2
RL BgIII N fragment.";
RL J. Virol. 49:724-730(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=333;
RX MEDLINE=84057718; PubMed=6315408;
RA McLauchlan J., Clements J.B.;
RT "DNA sequence homology between two co-linear loci on the HSV genome
RN which have different transforming abilities.";
RN EMBO J. 2:1953-1961(1983).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=HG52;
RA Dolan A.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases
CC -!- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
CC thioresoxin + H(2)O = ribonucleoside diphosphate + reduced
CC thioresoxin.
CC -!- COFACTOR: BINDS 2 IRON IONS.
CC -!- PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.
CC -!- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL CHAIN.
CC -!- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
CC SMALL CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M12700; AAA45807.1; -.
DR EMBL; X00048; CAA24930.1; -.
DR EMBL; Z86099; CAB06726.1; -.
DR PIR; A00528; WMBE32.
DR PIR; A00529; WMBE32.
DR InterPro: IPR000358; Ribonuc_redctse.
DR Pfam; PF00268; ribonuc_red_sm; 1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
DR Oxidoreductase; DNA replication; Iron.
KW METAL 91 91 IRON 1 (BY SIMILARITY).
FT METAL 121 121 IRON 1 AND 2 (BY SIMILARITY).
FT METAL 124 124 IRON 1 (BY SIMILARITY).
FT METAL 184 184 IRON 1 (BY SIMILARITY).
FT METAL 218 218 IRON 2 (BY SIMILARITY).
FT METAL 221 221 IRON 2 (BY SIMILARITY).
FT ACT_SITE 128 128 BY SIMILARITY.
FT CONFLICT 173 173 I -> V (IN REF. 2).
FT CONFLICT 235 235 G -> D (IN REF. 2).
SQ SEQUENCE 337 AA; 37625 MW; 1E27E9599EE2120 CRC64;

Query Match 75.0%; Score 27; DB 1; Length 337;
Best Local Similarity 62.5%; Pred. No. 43;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
| | | | |
Db 200 TNNLLRVT 207

RESULT 9
RIR2_HSV11
ID RIR2_HSV11 STANDARD; PRT; 340 AA.

```

```

AC P10224;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-AUG-1990 (Rel. 15, Last annotation update)
DE Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)
DE (Ribonucleotide reductase) (38 kDa subunit).
GN UL40.
OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88274327; PubMed=2839594;
RA McGeoch D.J., Dairymple M.A., Davison A.J., Dolan A., Frame M.C.,
RA McNeab D., Perry I.J., Scott J.E., Taylor P.;
RT "The complete DNA sequence of the long unique region in the genome of
RN herpes simplex virus type 1.";
RL J. Gen. Virol. 69:1531-1574(1988).
CC -!- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
CC thioresoxin + H(2)O = ribonucleoside diphosphate + reduced
CC thioresoxin.
CC -!- COFACTOR: BINDS 2 IRON IONS.
CC -!- PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.
CC -!- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL CHAIN.
CC -!- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
CC SMALL CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D10879; BAA01686.1; -.
DR EMBL; X14112; CAA32303.1; -.
DR PIR; D30088; WMBE57.
DR InterPro: IPR000358; Ribonuc_redctse.
DR Pfam; PF00268; ribonuc_red_sm; 1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
DR Oxidoreductase; DNA replication; Iron.
KW METAL 94 94 IRON 1 (BY SIMILARITY).
FT METAL 124 124 IRON 1 AND 2 (BY SIMILARITY).
FT METAL 127 127 IRON 1 (BY SIMILARITY).
FT METAL 187 187 IRON 2 (BY SIMILARITY).
FT METAL 221 221 IRON 2 (BY SIMILARITY).
FT METAL 224 224 IRON 2 (BY SIMILARITY).
FT ACT_SITE 131 131 BY SIMILARITY.
SQ SEQUENCE 340 AA; 38019 MW; 4B4ED994BF74FD3F CRC64;

Query Match 75.0%; Score 27; DB 1; Length 340;
Best Local Similarity 62.5%; Pred. No. 44;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
| | | | |
Db 203 TNNLLRVT 210

RESULT 10
RIR2_HSV1K
ID RIR2_HSV1K STANDARD; PRT; 340 AA.
AC P06474;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)
DE (Ribonucleotide reductase) (38 kDa subunit).
GN UL40.
OS Herpes simplex virus (type 1 / strain KOS).

```

OC Viruses: dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Simplexvirus.  
 OX NCBI\_TaxID=10306;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=6292456;  
 RX Draper K.G., Frink R.J., Wagner E.K.;  
 RA "Detailed characterization of an apparently unspliced beta herpes  
 RT simplex virus type 1 gene mapping in the interior of another.";  
 RL J. Virol. 43:1123-1128(1982).  
 RN [2]  
 RN REVISIONS.  
 RP Wagner E.K.;  
 RA Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
 RL  
 CC -!- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.  
 CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized  
 CC thiorodoxin + H(2)O = ribonucleoside diphosphate + reduced  
 CC thiorodoxin.  
 CC  
 CC -!- COFACTOR: BINDS 2 IRON IONS.  
 CC -!- PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.  
 CC -!- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL CHAIN.  
 CC -!- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE  
 CC SMALL CHAIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: J02212; AAA66436.1; -;  
 CC InterPro: IPR000358; Ribonuc\_redctse.  
 CC Pfam: PF00268; ribonuc\_red\_sm; 1.  
 CC PROSITE: PS00368; RIBRED\_SMALL; 1.  
 CC Oxidoreductase; DNA replication; Iron.  
 KW METAL 94 94 IRON 1 (BY SIMILARITY).  
 FT METAL 124 124 IRON 1 AND 2 (BY SIMILARITY).  
 FT METAL 127 127 IRON 1 (BY SIMILARITY).  
 FT ACT\_SITE 131 131 BY SIMILARITY.  
 FT ACT\_SITE 131 131 BY SIMILARITY.  
 SQ SEQUENCE 340 AA; 37966 MW; 921DC04B9D278DE5 CRC64;  
 Query Match 75.0%; Score 27; DB 1; Length 340;  
 Best Local Similarity 62.5%; Pred. No. 44;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 TNNVLQXT 8  
 DB 203 TNNLLRVT 210  
 RESULT 11  
 SOXB\_RHOSO  
 ID SOXB\_RHOSO STANDARD; PRT; 365 AA.  
 AC P54997;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Dibenzoethiophene desulfurization enzyme B (EC 3.1.2.24).  
 GN SOXB OR DSZB.  
 OS Rhodococcus sp. (strain IGTS8).  
 OC Plasmid.  
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.  
 OX NCBI\_TaxID=1831;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=9505232; PubMed=7961424;  
 RX Denome S.A., Oldfield C., Nash L.J., Young K.D.;  
 RA "Characterization of the desulfurization genes from Rhodococcus sp.  
 RT strain IGTS8.";  
 RL J. Bacteriol. 176:6707-6716(1994).

RN [2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=96031556; PubMed=7574582;  
 RX Piddington C.S., Kovacevich B.R., Rambossek J.;  
 RA "Sequence and molecular characterization of a DNA region encoding the  
 RT dibenzothiophene desulfurization operon of Rhodococcus sp. strain  
 RT IGTS8.";  
 RL Appl. Environ. Microbiol. 61:468-475(1995).  
 CC -!- FUNCTION: PART OF A PATHWAY TO REMOVE COVALENTLY BOUND SULFUR FROM  
 CC DIBENZOTHIOPHENE (DBT) WITHOUT BREAKING CARBON-CARBON BONDS. THIS  
 CC ENZYME METABOLIZES DBT-SULFONE (DBTO2 OR DBT 5,5-DIOXIDE) TO 2-  
 CC HYDROXYBIPHENYL (2-HBP).  
 CC -!- CATALYTIC ACTIVITY: 2-(2-hydroxyphenyl)benzenesulfinate + H(2)O =  
 CC 2-hydroxybiphenyl + sulfite.  
 CC -!- COFACTOR: FMN (POTENTIAL).  
 CC -!- PATHWAY: SECOND STEP IN PATHWAY FROM DBT TO 2-HBP.  
 CC -!- SUBUNIT: HETERODIMER OF TWO SUBUNITS, SOXA AND SOXB.  
 CC -!- SIMILARITY: BELONGS TO THE NTAA/SNAA/DSZA(SOXA) FAMILY OF  
 CC MONOOXYGENASES.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: U08850; AAA56672.1; -;  
 CC EMBL: L37363; AAA99483.1; -;  
 CC Hydrolase; Monooxygenase; Flavoprotein; FMN; Plasmid.  
 KW SEQUENCE 365 AA; 39044 MW; DA6A867756DA23D6 CRC64;  
 SQ  
 Query Match 75.0%; Score 27; DB 1; Length 365;  
 Best Local Similarity 83.3%; Pred. No. 47;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TNNVLQ 6  
 DB 338 TNNLLQ 343  
 RESULT 12  
 CYOA\_HAEIN  
 ID CYOA\_HAEIN STANDARD; PRT; 521 AA.  
 AC P45021;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Probable cytochrome oxidase subunit I (EC 1.10.3.-).  
 GN H11076.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Haemophilus.  
 OX NCBI\_TaxID=727;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-Rd / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RA "Whole-genome random sequencing and assembly of Haemophilus influenzae  
 RT Rd.";  
 RL Science 269:496-512(1995).  
 CC -!- FUNCTION: PROBABLE CYTOCHROME OXIDASE SUBUNIT.  
 CC -!- COFACTOR: CONTAINS THE PROTOHEME IX CENTER B558 (BY SIMILARITY).

```

CC -!- SUBUNIT: HETERODIMER OF SUBUNITS I AND II (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (BY SIMILARITY).
CC -!- SIMILARITY: STRONG, TO BOTH E.COLI APPC AND CYDA. PROBABLE
CC ORTHOLOG OF THE ANCESTOR OF APPC/CYDA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32787; AAC22732.1; -
DR TIGR; H11076; -
DR InterPro; IPR002585; Bac_Ubq_Cox.
DR Pfam; PF01654; Bac_Ubq_Cox; 1.
KW Oxidoreductase; Electron transport; Transmembrane; Inner membrane;
KW Heme; Complete proteome.
FT DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 23 42 POTENTIAL.
FT DOMAIN 43 94 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 95 114 POTENTIAL.
FT DOMAIN 115 129 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 130 149 POTENTIAL.
FT DOMAIN 150 187 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 188 207 POTENTIAL.
FT DOMAIN 208 219 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 220 239 POTENTIAL.
FT DOMAIN 240 397 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 398 417 POTENTIAL.
FT DOMAIN 418 475 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 476 495 POTENTIAL.
FT DOMAIN 496 521 PERIPLASMIC (POTENTIAL).
FT METAL 186 186 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 521 AA; 58136 MW; FD44C20457D2265 CRC64;

Query Match 75.0%; Score 27; DB 1; Length 521;
Best Local Similarity 62.5%; Pred. No. 72;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
Db 366 TNNVVDAT 373
||||:|

RESULT 13
60IM_COXBU STANDARD; PRT; 638 AA.
AC P45650;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 60 kDa inner-membrane protein homolog.
OS Coxiella burnetii.
OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
OC Coxiella group; Coxiella.
OC NCBI_TaxID=777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nine Mile phase I / Bratislava;
RX MEDLINE=94350801; PubMed=8071197;
RA Suhan M., Chen S.Y., Thompson H.A., Hoover T.A., Hill A.,
RA Williams J.C.;
RT "Cloning and characterization of an autonomous replication sequence
RT from Coxiella burnetii.";
RL J. Bacteriol. 176:5233-5243(1994).
RN [2]
RP CONCEPTUAL TRANSLATION.
RA Bairoch A.;
RL Unpublished observations (AUG-1995).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane

```

```

CC (Probable).
CC -!- SIMILARITY: BELONGS TO THE OXAL / 60 KDA IMP FAMILY.
CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A FRAMESHIFT WAS
CC INTRODUCED IN POSITION 482. THERE IS PROBABLY ANOTHER ERROR
CC IN THE C-TERMINAL PART AND THE REAL SEQUENCE COULD BE SHORTER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U10529; AAA56919.1; ALT FRAME.
DR EMBL; U10529; AAA56921.1; ALT FRAME.
DR InterPro; IPR001708; 60kDa_innermemb.
DR Pfam; PF02096; 60KD_IMP; 1.
DR PRINTS; PR00701; 60KDINNERMP.
KW Transmembrane; Inner membrane.
FT TRANSMEM 3 23 POTENTIAL.
FT TRANSMEM 346 366 POTENTIAL.
FT TRANSMEM 369 389 POTENTIAL.
FT TRANSMEM 436 456 POTENTIAL.
SQ SEQUENCE 638 AA; 72903 MW; 6E670FC9EA2E80E2 CRC64;

Query Match 75.0%; Score 27; DB 1; Length 638;
Best Local Similarity 83.3%; Pred. No. 92;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
Db 533 TNNVVQ 538
||||:|

RESULT 14
COAT_ADVQ STANDARD; PRT; 647 AA.
AC P24029;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein VP1 [Contains: Coat protein VP2].
OS Aleutian mink disease parvovirus (strain G) (ADV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OC NCBI_TaxID=10783;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88275062; PubMed=2839709;
RA Bloom M.E., Alexandersen S., Perryman S., Lechner D.,
RA Wolfenbarger J.B.;
RT "Nucleotide sequence and genomic organization of Aleutian mink
RT disease parvovirus (ADV): sequence comparisons between a
RT nonpathogenic and a pathogenic strain of ADV.";
RL J. Virol. 62:2903-2915(1988).
CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M20036; AAA66615.1; -
DR PIR; A35529; VCPVPAP.
DR PIR; B36760; B36760.
DR HSSP; P30129; 4DPV.
DR InterPro; IPR001403; parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein; Glycoprotein.
DOMAIN 22 39 POLY-GLY.

```

FT CARBOHYD 16 16 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 399 399 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 647 AA; 73517 MW; 22CE812094FFBFA CRC64;

Query Match 75.0%; Score 27; DB 1; Length 647;

Best Local Similarity 62.5%; Pred. No. 94;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8

Db 185 TNNILPYT 192

## RESULT 15

NTPI\_MSEPV STANDARD; PRT; 647 AA.  
AC Q9YV39;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Nucleoside triphosphatase I (EC 3.6.1.15) (Nucleoside triphosphate  
phosphohydrolase I) (NPH I).  
GN MSV053.  
OS Melanoplus sanguinipes entomopoxvirus (MsePV).  
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;  
OC Entomopoxvirus B.  
OX NCBI\_TaxID=83191;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Tucson;  
RX MEDLINE=99102612; PubMed=9847359;  
RA Afonso C.L., Tulman E.R., Lu Z., Oma E., Kutish G.F., Rock D.L.;  
RT "The genome of Melanoplus sanguinipes entomopoxvirus.";  
RL J. Virol. 73:533-552(1999).  
CC -!- FUNCTION: SERVES TWO ROLES IN TRANSCRIPTION; IT ACTS IN CONCERT  
WITH VIRAL TERMINATION FACTOR/CAPPING ENZYME TO CATALYZE RELEASE  
OF UUUUUU-CONTAINING NASCENT RNA FROM THE ELONGATION COMPLEX, AND  
IT ACTS BY ITSELF AS A POLYMERASE ELONGATION FACTOR TO FACILITATE  
READTHROUGH OF INTRINSIC PAUSE SITES (BY SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: NTP + H(2)O = NDP + phosphate.  
CC -!- SIMILARITY: BELONGS TO THE NPH I SUBFAMILY OF HELICASES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF063866; AAC37824.1; -;  
DR InterPro; IPR001650; Helicase\_C.  
DR InterPro; IPR000330; SNF2\_N.  
DR Pfam; PF00176; SNF2\_N; 1.  
DR Pfam; PF00271; Helicase\_C; 1.  
DR SMART; SM00490; HELICCC; 1.  
DR HydroLase; ATP-binding; Transcription.  
FT DOMAIN 40 322 SNF2\_N.  
FT DOMAIN 404 490 HELICASE\_C.  
FT NP\_BIND 61 68 ATP (BY SIMILARITY).  
FT SITE 150 153 DEXH BOX.  
SQ SEQUENCE 647 AA; 75134 MW; E78C7F768235D51D CRC64;

Query Match 75.0%; Score 27; DB 1; Length 647;  
Best Local Similarity 83.3%; Pred. No. 94;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
Db 197 TNNVLE 202

Search completed: February 26, 2003, 15:37:37  
Job time : 30 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:33:36 ; Search time 29 Seconds  
(without alignments)  
56.841 Million cell updates/sec

Title: US-09-476-485A-24  
Perfect score: 36  
Sequence: 1 TNNVLQXT 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

#### Database :

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriaph.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	34	94.4	272	10 Q9ZTA9	Q9ZTA9 Dolichos la
2	34	94.4	279	10 Q9W7M4	Q9W7M4 phaseolus v
3	31	86.1	79	10 Q94AJ7	Q94AJ7 arabidopsis
4	30	83.3	115	16 Q26052	Q26052 helicobacte
5	30	83.3	115	16 Q9ZJA0	Q9ZJA0 helicobacte
6	30	83.3	248	5 Q9XW20	Q9XW20 caenorhabdi
7	30	83.3	345	16 Q97F80	Q97F80 clostridium
8	30	83.3	541	10 Q94LJ3	Q94LJ3 zea mays (m
9	30	83.3	544	16 Q34609	Q34609 neisseria m
10	30	83.3	822	12 Q39989	Q39989 human herpe
11	30	83.3	822	12 Q56280	Q56280 human herpe
12	30	83.3	822	12 Q96910	Q96910 human herpe
13	30	83.3	960	10 Q82318	Q82318 arabidopsis
14	30	83.3	1268	2 Q9RER7	Q9RER7 bacillus sp
15	30	83.3	2724	4 Q96SE4	Q96SE4 homo sapien
16	29	80.6	170	5 Q9NBF9	Q9NBF9 philodina r

17	29	80.6	284	10 Q9FY09	Q9FY09 sophora fla
18	29	80.6	311	16 Q9ABM7	Q9ABM7 caulobacter
19	29	80.6	374	12 Q66210	Q66210 paramecium
20	29	80.6	374	12 Q98583	Q98583 paramecium
21	29	80.6	393	10 Q9ZSC0	Q9ZSC0 arabidopsis
22	29	80.6	407	10 Q9T0C1	Q9T0C1 arabidopsis
23	29	80.6	501	5 Q17319	Q17319 ceratitis c
24	29	80.6	524	10 Q9SKG4	Q9SKG4 arabidopsis
25	29	80.6	613	2 Q93LN4	Q93LN4 xanthomonas
26	29	80.6	649	12 Q91GN0	Q91GN0 epiphyas po
27	29	80.6	674	17 Q96XG2	Q96XG2 sulfolobus
28	29	80.6	677	16 Q9CCP2	Q9CCP2 mycobacteri
29	28	77.8	167	13 Q9IBS8	Q9IBS8 xenopus lae
30	28	77.8	163	2 Q53297	Q53297 escherichia
31	28	77.8	204	13 Q91P99	Q91P99 xenopus lae
32	28	77.8	212	15 Q9WGU7	Q9WGU7 human immun
33	28	77.8	351	5 Q966A1	Q966A1 caenorhabdi
34	28	77.8	363	16 Q8XRT3	Q8XRT3 raistonia s
35	28	77.8	390	16 Q8Y2H4	Q8Y2H4 raistonia s
36	28	77.8	405	16 Q9ZDH0	Q9ZDH0 rickettsia
37	28	77.8	411	16 Q92ID4	Q92ID4 rickettsia
38	28	77.8	466	5 Q22395	Q22395 caenorhabdi
39	28	77.8	554	3 Q08887	Q08887 saccharomyc
40	27	75.0	134	10 Q65553	Q65553 arabidopsis
41	27	75.0	186	3 Q9P7X0	Q9P7X0 schizosacch
42	27	75.0	200	2 Q8RTM2	Q8RTM2 helicobacte
43	27	75.0	208	16 Q92I77	Q92I77 rickettsia
44	27	75.0	229	12 Q9QC4	Q9QC4 aleutian mi
45	27	75.0	230	12 Q9QC5	Q9QC5 aleutian mi

#### ALIGNMENTS

#### RESULT 1

Q9ZTA9 ID Q9ZTA9 PRELIMINARY; PRT; 272 AA.  
AC Q9ZTA9;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Mannose lectin.  
GN FRIL.  
OS Dolichos lab lab (Field bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Dolichos.  
OX NCBI\_TaxID=35936;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=COTYLEDON;  
RX MEDLINE=99110944; PubMed=9892687;  
RA Colucci G., Moore J.G., Feldman M., Chrispeels M.J.;  
RT "cDNA cloning of FRIL, a lectin from Dolichos lablab, that preserves  
RT hemagglutinating activity in suspension culture."  
RL Proc. Natl. Acad. Sci. U.S.A. 96:646-650(1999).  
DR EMBL; AF067417; AAD10734.1; -;  
DR HSSP; P02866; 1ONA.  
DR InterPro; IPR000985; Lectin\_lega.  
DR InterPro; IPR001220; Lectin\_legB.  
DR Pfam; PF00138; lectin\_legA; 1.  
DR Pfam; PF00139; lectin\_legB; 1.  
DR ProDom; PD000671; lectin\_legA; 1.  
DR ProDom; PD000711; lectin\_legB; 1.  
DR PROSITE; PS00308; LECTIN\_LEGUME\_ALPHA; 1.  
DR PROSITE; PS00307; LECTIN\_LEGUME\_BETA; UNKNOWN; 1.  
SQ SEQUENCE 272 AA; 29900 MW; EA6C004307441495 CRC64;

Query Match 94.4%; Score 34; DB 10; Length 272;  
Best Local Similarity 87.5%; Pred.No. 15;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 TNNVLOXT 8
DB 35 TNNVLOXT 42

RESULT 2
Q9M7M4
ID Q9M7M4 PRELIMINARY; PRT; 279 AA.
AC Q9M7M4;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-DEC-2001 (TRENBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Mannose lectin FRL (Fragment).
OS Phaseolus vulgaris (kidney bean) (French bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
OX NCBI_TaxID=3885;
RN [1]
RN SEQUENCE FROM N.A.
RP Moore J.G., Colucci G., Fuchs C.A., Hicklin D.J., Chrispeels M.J.,
RA Feldman M.;
RT "A new lectin in red kidney bean called PVFRL stimulates
RT proliferation of NIH3T3 cells expressing the Flt3 receptor.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF121458; AAF28739.1;
DR HSSP: P02866; IONA.
DR InterPro: IPR000985; Lectin_legA.
DR InterPro: IPR001220; Lectin_legB.
DR Pfam: PF00138; lectin_legA; 1.
DR Pfam: PF00139; lectin_legB; 1.
DR ProDom: PD000671; Lectin_legA; 1.
DR ProDom: PD000711; Lectin_legB; 1.
DR PROSITE: PS00308; LECTIN_LEGUME_ALPHA; 1.
DR PROSITE: PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
KW Lectin.
FT NON_TER
FT SEQUENCE 279 AA; 31102 MW; F8919CF8B3E4652 CRC64;

Query Match 94.4%; Score 34; DB 10; Length 279;
Best Local Similarity 87.5%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
DB 27 TNNVLOXT 34

RESULT 3
Q94AJ7
ID Q94AJ7 PRELIMINARY; PRT; 79 AA.
AC Q94AJ7;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Hypothetical 8.7 kDa protein.
GN AT4G33665 OR AT4G33666.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RN SEQUENCE FROM N.A.
RP Yamada K., Liu S.X., Pham P.K., Banh J., Dale J.M., Goldsmith A.D.,
RA Jiang P.X., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kaniya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
RA Ecker J.R., Theologis A.;
RT "Full length cDNA of gene AT4G33665 (not previously annotated).";

Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
RP Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Etgu P., Lee J.M.,
RA Toriumi M., Yu G., Brooks S., Chao Q., Chen H., Karlin-Neumann G.,
RA Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shinn P.,
RA Southwick A., Davis R.W., Ecker J.R., Theologis A.;
RT "Arabidopsis Open Reading Frame (ORF) Clones.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY045998; AAK76672.1;
DR EMBL: AY079343; AAL85074.1;
KW Hypothetical protein.
SQ SEQUENCE 79 AA; 8660 MW; F8187D6D858360D0 CRC64;

Query Match 86.1%; Score 31; DB 10; Length 79;
Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
DB 20 TNNVLOXT 27

RESULT 4
O26052
ID O26052 PRELIMINARY; PRT; 115 AA.
AC O26052;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Hypothetical protein HPI524.
GN HPI524.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547(1997).
DR EMBL: AE000650; AAD08566.1;
DR TIGR: HPI524;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 115 AA; 13287 MW; 1B56AA20E27EBE9D CRC64;

Query Match 83.3%; Score 30; DB 16; Length 115;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLO 6
DB 93 TNNVLO 98

RESULT 5
Q9ZJAO
ID Q9ZJAO PRELIMINARY; PRT; 115 AA.
AC Q9ZJAO;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)

```



DE PUTATIVE.  
GN JHP1413.  
OS Helicobacter pylori J99 (Campylobacter pylori J99).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=85963;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99120557; PubMed=9923682;  
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,  
RA Smith D.R., Noonan B., Gull B.C., deJonge B.L., Carmel G.,  
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,  
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,  
RA Trust T.J.;  
RT "Genomic sequence comparison of two unrelated isolates of the human  
RT gastric pathogen Helicobacter pylori."  
RL Nature 397:176-180(1999).  
DR EMBL: AE001564; AAD06992.1; -.  
KW Complete proteome.  
SQ SEQUENCE 115 AA; 13299 MW; CD07B03053966DBF CRC64;

Query Match 83.3%; Score 30; DB 16; Length 115;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
| | | | |  
DB 93 TNNVLQ 98

RESULT 6  
Q9XW20  
ID Q9XW20 PRELIMINARY; PRT; 248 AA.  
AC Q9XW20;  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Y18D10A.11 protein.  
GN Y18D10A.11.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Harris B.R.;  
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=98519116;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
RT investigating biology."  
RL Science 282:2012-2018(1998).  
DR EMBL: AL034393; CAA22314.1; -.  
SQ SEQUENCE 248 AA; 27963 MW; E7COA295197C6F6E CRC64;

Query Match 83.3%; Score 30; DB 5; Length 248;  
Best Local Similarity 85.7%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNVLQXT 8  
| | | | |  
DB 150 NNVLQAT 156

RESULT 7  
Q97F80  
ID Q97F80 PRELIMINARY; PRT; 345 AA.  
AC Q97F80;  
DT 01-OCT-2001 (TREMBlrel. 18, Created)  
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)  
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)

DE Sporulation protein spoIID.  
GN CAC2861.  
OS Clostridium acetobutylicum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
OC Clostridiales; Clostridiaceae; Clostridium.  
OX NCBI\_TaxID=1488;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=ATCC 824 / DSM 792 / VKM B-1787;  
RA MEDLINE=21359325; PubMed=11466286;  
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,  
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hittl J., Wolf Y.I.,  
RA Tatusov R.L., Sabath E.V., Doucette-Stamm L., Soucaille P., Daly M.J.,  
RA Bennett G.N., Koonin E.V., Smith D.R.;  
RT "Genome sequence and comparative analysis of the solvent-producing  
RT bacterium Clostridium acetobutylicum."  
RL J. Bacteriol. 183:4823-4838(2001).  
DR EMBL: AE007784; AAK80804.1; -.  
KW Complete proteome.  
SQ SEQUENCE 345 AA; 38653 MW; 2082986E5BDDFC3D CRC64;

Query Match 83.3%; Score 30; DB 16; Length 345;  
Best Local Similarity 85.7%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNVLQXT 8  
| | | | |  
DB 271 NNVLQAT 277

RESULT 8  
Q94IJ3  
ID Q94IJ3 PRELIMINARY; PRT; 541 AA.  
AC Q94IJ3;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Somatic embryogenesis receptor-like kinase 3 (Fragment).  
GN SERK3  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. A188;  
RA Baudino S., Hansen S., Brettschneider R., Hecht V.F., Dresselhaus T.,  
RA Loerz H., Dumas C., Rogowsky P.M.;  
RT "Molecular characterisation of two novel maize LRR receptor-like  
RT kinases, which belong to the SERK gene family."  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ400870; CAC37642.1; -.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR001611; LRR.  
DR InterPro: IPR002290; Ser\_thr\_pkinase.  
DR Pfam: PF00560; LRR; 5.  
DR Pfam: PF00069; pkinase; 1.  
DR ProDom: PD000001; Euk\_pkinase; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; UNKNOWN\_1.  
KW Kinase; Receptor.  
FT NON\_TER 1 541  
FT NON\_TER 1 541  
SQ SEQUENCE 541 AA; 59158 MW; FCC6BEA86BDF8587 CRC64;

Query Match 83.3%; Score 30; DB 10; Length 541;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
| | | | |





OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Abbas A.R., Linn S.M.;  
 RT "Homo sapiens polymerase (DNA-directed), theta (POLQ), mRNA."  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY032677; AAK39635.1; -  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR001098; DNA\_pol.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR Pfam; PF00270; DEAD; 1.  
 DR Pfam; PF00476; DNA\_pol\_A; 1.  
 DR Pfam; PF00271; Helicase\_C; 1.  
 DR PROSITE; PS00447; DNA\_POLYMERASE\_A; UNKNOWN\_1.  
 KW ATP-binding; Helicase.  
 SQ SEQUENCE 2724 AA; 305148 MW; 30B88663614E65DC CRC64;

Query Match 83.38; Score 30; DB 4; Length 2724;  
 Best Local Similarity 100.0%; Pred. No. 9.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
 |||||  
 Db 1896 TNNVLQ 1901

Search completed: February 26, 2003, 15:38:13  
 Job time : 31 secs



PT cells - particularly haematopoietic progenitors, useful for bone  
 PT marrow reconstitution after ablative therapy, and to increase DNA  
 PT transfer in gene therapy

XX Claim 1; Page 46; 72pp; English.

XX The present sequence represents a peptide of lectin derived progenitor  
 CC a cell preservation factor. The protein is used to preserve unipotent,  
 CC pluripotent or totipotent progenitor cells, especially haematopoietic  
 CC cells, and also progenitors from nerve, muscle, skin, gut, bone,  
 CC kidney, liver, pancreas or thymus. Specific applications are  
 CC preservation of cultured cells intended for administration after  
 CC (anticancer) myeloablative therapy (bone marrow or whole-body irradiation  
 CC or chemotherapy) to reconstitute the haematopoietic system; enrichment  
 CC of progenitor cells (e.g. during ex vivo purging of malignant cells);  
 CC treatment of tissues containing haematopoietic progenitors for subsequent  
 CC transplant to improve haematopoietic competence; improving transfer of  
 CC exogenous DNA to progenitor cells (in gene therapy of various  
 CC haematological disorders, e.g. sickle-cell anaemia); and protection  
 CC against ablative therapy (to eliminate proliferating cells specifically),  
 CC followed by re-establishment of differentiation and proliferation of  
 CC preserved progenitors. The protein, when linked to magnetic beads, may  
 CC also be used to isolate cells that express the FLK2/FLT3 receptor.

XX Sequence 8 AA;

Query Match 94.4%; Score 34; DB 20; Length 8;  
 Best Local Similarity 87.5%; Pred. No. 7.8e+05;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8  
 | | | | |  
 Db 1 TNNVLQXT 8

# RESULT 2

AAAG62889  
 ID AAAG62889 standard; peptide; 8 AA.

XX  
 AC AAAG62889;

XX 17-SEP-2001 (first entry)

XX Peptide derived from a hyacinth bean FRIL polypeptide.

XX FRIL; FLK2/FLT3 tyrosine kinase receptor-interacting lectin;  
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;  
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
 KW severe combined immunodeficiency; aplastic anemia; tissue repair.

OS Dolichos lab lab.

XX Key Location/Qualifiers

FT Misc-difference 7 /note= "not specified"

XX WO200149851-A1.

PN 12-JUL-2001.

XX 30-DEC-1999; 99WO-US31307.

XX 30-DEC-1999; 99WO-US31307.

XX (PHYL-) PHYLOGIX LLC.

XX Colucci MG, Chrispeels MJ, Moore JG;

XX WPI; 2001-441882/47.

XX Legume Progenitor cell preservation factors for in vivo or ex vivo  
 PT preservation of haematopoietic progenitor cells and as therapeutics for  
 PT alleviating/reducing progenitor cell-depleting activity of cancer

## therapeutics -

XX Disclosure; Page 19; 173pp; English.

XX The present sequence is derived from a FRIL (FLK2/FLT3 tyrosine kinase  
 CC receptor-interacting lectin) polypeptide. The specification describes a  
 CC composition of one or more members of FRIL family of progenitor cell  
 CC preservation factors. The composition is useful for alleviating or  
 CC reducing the hematopoietic progenitor cell-depleting activity of a  
 CC therapeutic treatment, including radiotherapeutic and/or  
 CC chemotherapeutic treatments. Administration of FRIL compositions to a  
 CC patient prior to treatment of the patient with a therapeutic treatment  
 CC having a hematopoietic progenitor cell-depleting activity alleviates or  
 CC reduces the hematopoietic progenitor cell-depleting activity of the  
 CC therapeutic treatment in the patient. FRIL family members are useful for  
 CC isolating population of progenitor cells, hemangioblasts, and mesenchymal  
 CC stem cells. The composition is administered to reduce progenitor cell  
 CC depleting effects of chemotherapeutics, so that the patient can receive  
 CC a higher dose of the chemotherapeutic and preferably recover from cancer.  
 CC It is also administered to patients having, or predisposed to developing  
 CC a condition where the patients hematopoietic progenitor cells are  
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
 CC The isolated mesenchymal cells are useful for tissue repair.

XX Sequence 8 AA;

Query Match 94.4%; Score 34; DB 22; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8  
 | | | | |  
 Db 1 TNNVLQXT 8

# RESULT 3

AAW87973

ID AAW87973 standard; Protein; 264 AA.

XX  
 AC AAW87973;

XX 13-APR-1999 (first entry)

XX A lectin derived progenitor cell preservation factor.

XX Lectin derived progenitor cell preservation factor; progenitor cell;  
 KW haematopoietic cell; cultured cell preservation; anticancer therapy;  
 KW myeloablative therapy; sickle-cell anaemia; ablative therapy protection;  
 KW FLK2/FLT3 receptor.

OS Dolichos lab lab.

XX WO9859038-A1.

XX 30-DEC-1998.

XX 23-JUN-1998; 98WO-US13046.

XX 24-JUN-1997; 97US-0881189.

XX (IMCL-) IMCLONE SYSTEMS INC.  
 PA (REGC) UNIV CALIFORNIA.

XX Chrispeels MJ, Colucci MG, Moore JG;  
 XX WPI; 1999-081274/07.  
 DR N-PSDB; AAX03593.

XX New nucleic acid encoding plant lectin that preserves progenitor  
 PT cells - particularly haematopoietic progenitors, useful for bone  
 PT marrow reconstitution after ablative therapy, and to increase DNA  
 PT transfer in gene therapy

Claim 1; Page 30-31; 72pp; English.

The present sequence represents a lectin derived progenitor cell preservation factor. The protein is used to preserve unipotent, pluripotent or totipotent progenitor cells, especially haematopoietic cells, and also progenitors from nerve, muscle, skin, gut, bone, kidney, liver, pancreas or thymus. Specific applications are preservation of cultured cells intended for administration after (anticancer) myeloablative therapy (bone marrow or whole-body irradiation or chemotherapy) to reconstitute the haematopoietic system; enrichment of progenitor cells (e.g. during ex vivo purging of malignant cells); treatment of tissues containing haematopoietic progenitors for subsequent transplant to improve haematopoietic competence; improving transfer of exogenous DNA to progenitor cells (in gene therapy of various haematological disorders, e.g. sickle-cell anaemia); and protection against ablative therapy (to eliminate proliferating cells specifically), followed by re-establishment of differentiation and proliferation of preserved progenitors. The protein, when linked to magnetic beads, may also be used to isolate cells that express the FLK2/Flt3 receptor.

SQ Sequence 264 AA;

Query Match 94.4%; Score 34; DB 20; Length 264;  
Best Local Similarity 87.5%; Pred. No. 9.1;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TNNVLQXT 8  
|||||  
Db 27 TNNVLQVT 34

RESULT 4

AAG62890  
ID AAG62890 standard; Protein; 264 AA.

AC AAG62890;

DT 17-SEP-2001 (first entry)

DE Amino acid sequence of a hyacinth bean FRIL polypeptide.

XX FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;  
KW progenitor cell preservation factor; radiotherapy; chemotherapy;  
KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
XX severe combined immunodeficiency; aplastic anemia; tissue repair.

OS Dolichos lab lab.

XX WO200149851-A1.

PN 12-JUL-2001.

PD 30-DEC-1999; 99WO-US31307.

PF 30-DEC-1999; 99WO-US31307.

PR (PHYL-) PHYLOGIX LLC.

PI Colucci MG, Chrispeels MJ, Moore JG;

XX WPI; 2001-441882/47.

DR N-PSDB; AAH42287.

PT Legume Progenitor cell preservation factors for in vivo or ex vivo preservation of hematopoietic progenitor cells and as therapeutics for alleviating/reducing progenitor cell-depleting activity of cancer therapeutics

XX Example 1; Page 54-55; 173pp; English.

XX The present sequence represents a FRIL (Flk2/Flt3 tyrosine kinase receptor-interacting lectin) polypeptide. The specification describes a composition of one or more members of FRIL family of progenitor cell

CC

CC preserving factors. The composition is useful for alleviating or reducing the hematopoietic progenitor cell-depleting activity of a therapeutic treatment, including radiotherapeutic and/or chemotherapeutic treatments. Administration of FRIL compositions to a patient prior to treatment of the patient with a therapeutic treatment having a hematopoietic progenitor cell-depleting activity alleviates or reduces the hematopoietic progenitor cell-depleting activity of the therapeutic treatment in the patient. FRIL family members are useful for isolating population of progenitor cells, hemangioblasts, and mesenchymal stem cells. The composition is administered to reduce progenitor cell depleting effects of chemotherapeutics, so that the patient can receive a higher dose of the chemotherapeutic and preferably recover from cancer. It is also administered to patients having, or predisposed to developing a condition where the patients hematopoietic progenitor cells are depleted, such as severe combined immunodeficiency or aplastic anemia. The isolated mesenchymal cells are useful for tissue repair.

XX Sequence 264 AA;

Query Match 94.4%; Score 34; DB 22; Length 264;  
Best Local Similarity 87.5%; Pred. No. 9.1;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TNNVLQXT 8

Db 27 TNNVLQVT 34

RESULT 5

AAG62894  
ID AAG62894 standard; Protein; 286 AA.

AC AAG62894;

DT 17-SEP-2001 (first entry)

DE Alpha-amylase inhibitor signal peptide and FRIL fusion.

XX FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;  
KW progenitor cell preservation factor; radiotherapy; chemotherapy;  
KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
XX severe combined immunodeficiency; aplastic anemia; tissue repair;  
XX alpha-amylase inhibitor gene.

OS Synthetic.

OS Unidentified.

OS Dolichos lab lab.

XX WO200149851-A1.

PN 12-JUL-2001.

PD 30-DEC-1999; 99WO-US31307.

PF 30-DEC-1999; 99WO-US31307.

PR (PHYL-) PHYLOGIX LLC.

PI Colucci MG, Chrispeels MJ, Moore JG;

XX WPI; 2001-441882/47.

DR N-PSDB; AAH42295.

PT Legume Progenitor cell preservation factors for in vivo or ex vivo preservation of hematopoietic progenitor cells and as therapeutics for alleviating/reducing progenitor cell-depleting activity of cancer therapeutics

XX Example 1; Page 59; 173pp; English.

XX The present sequence represents fusion protein of alpha-amylase inhibitor signal peptide and FRIL (Flk2/Flt3 tyrosine kinase receptor-interacting lectin). The specification describes a composition of one or more members

CC of FRIL family of progenitor cell preservation factors. The composition  
 CC is useful for alleviating or reducing the hematopoietic progenitor  
 CC cell-depleting activity of a therapeutic treatment, including  
 CC radiotherapeutic and/or chemotherapeutic treatment. Administration of  
 CC FRIL compositions to a patient prior to treatment of the patient with  
 CC a therapeutic treatment having a hematopoietic progenitor cell-depleting  
 CC activity alleviates or reduces the hematopoietic progenitor  
 CC cell-depleting activity of the therapeutic treatment in the patient.  
 CC FRIL family members are useful for isolating population of progenitor  
 CC cells, hemangioblasts, and mesenchymal stem cells. The composition is  
 CC administered to reduce progenitor cell depleting effects of  
 CC chemotherapeutic, so that the patient can receive a higher dose of the  
 CC chemotherapeutic and preferably recover from cancer. It is also  
 CC administered to patients having, or predisposed to developing a  
 CC condition where the patients hematopoietic progenitor cells are  
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
 CC The isolated mesenchymal cells are useful for tissue repair.

XX  
 XX  
 SQ Sequence 286 AA;  
 Query Match 94.4%; Score 34; DB 22; Length 286;  
 Best Local Similarity 87.5%; Pred No. 10;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8  
 |||||  
 Db 49 TNNVLQVT 56

RESULT 6  
 AAG62898  
 ID AAG62898 standard; Protein; 303 AA.

XX  
 AC AAG62898;

DT 17-SEP-2001 (first entry)

XX Amino acid sequence of a french bean FRIL polypeptide.

XX FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;  
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;  
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
 KW severe combined immunodeficiency; aplastic anemia; tissue repair.

XX Phaseolus vulgaris.

XX WO200149851-A1.

XX 12-JUL-2001.

PF 30-DEC-1999; 99WO-US31307.

XX 30-DEC-1999; 99WO-US31307.

XX (PHYL-) PHYLOGIX LLC.

XX Colucci MG, Chrispeels MJ, Moore JG;

XX WPI; 2001-441882/47.

DR N-PSDB; AAH42306.

XX Legume progenitor cell preservation factors for in vivo or ex vivo  
 PT preservation of hematopoietic progenitor cells and as therapeutics for  
 PT alleviating/reducing progenitor cell-depleting activity of cancer  
 PT therapeutics -

XX Example 5; Page 81; 173pp; English.

XX The present sequence represents a FRIL (Flk2/Flt3 tyrosine kinase  
 CC receptor-interacting lectin) polypeptide. The specification describes a  
 CC composition of one or more members of FRIL family of progenitor cell  
 CC preservation factors. The composition is useful for alleviating or  
 CC reducing the hematopoietic progenitor cell-depleting activity of

CC a therapeutic treatment, including radiotherapeutic and/or  
 CC chemotherapeutic treatments. Administration of FRIL compositions to a  
 CC patient prior to treatment of the patient with a therapeutic treatment  
 CC having a hematopoietic progenitor cell-depleting activity alleviates or  
 CC reduces the hematopoietic progenitor cell-depleting activity of the  
 CC therapeutic treatment in the patient. FRIL family members are useful for  
 CC isolating population of progenitor cells, hemangioblasts, and mesenchymal  
 CC stem cells. The composition is administered to reduce progenitor cell  
 CC depleting effects of the chemotherapeutic, so that the patient can receive  
 CC a higher dose of the chemotherapeutic and preferably recover from cancer.  
 CC It is also administered to patients having, or predisposed to developing  
 CC a condition where the patients hematopoietic progenitor cells are  
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
 CC The isolated mesenchymal cells are useful for tissue repair.

XX  
 XX  
 SQ Sequence 303 AA;

Query Match 94.4%; Score 34; DB 22; Length 303;  
 Best Local Similarity 87.5%; Pred. No. 11;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8  
 |||||  
 Db 27 TNNVLQLT 34

RESULT 7  
 AAG59708  
 ID AAG59708 standard; Protein; 73 AA.

XX  
 AC AAG59708;

XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 77260.

XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

XX Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 23-APR-1999; 99US-0130891.

XX 28-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 04-MAY-1999; 99US-0132407.

XX 05-MAY-1999; 99US-0132484.

XX 06-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 07-MAY-1999; 99US-0132487.

XX 11-MAY-1999; 99US-0132863.

XX 14-MAY-1999; 99US-0134256.

XX 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134219.



PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140595.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.

```
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 86.1%; Score 31; DB 21; Length 73;
Best Local Similarity 75.0%; Pred. No. 9.6;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TNNVLOXT 8
Db 14 TNNVLOXT 21

RESULT 8
AAG59707
ID AAG59707 standard; Protein; 79 AA.
XX AC AAG59707;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 77259.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EPI033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 28-APR-1999; 99US-0130891.
XX PR 30-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 30-APR-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132485.
XX PR 05-MAY-1999; 99US-0132486.
XX PR 06-MAY-1999; 99US-0132487.
XX PR 07-MAY-1999; 99US-0132863.
XX PR 11-MAY-1999; 99US-0134256.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134370.
XX PR 18-MAY-1999; 99US-0134768.
XX PR 19-MAY-1999; 99US-0134941.
XX PR 20-MAY-1999; 99US-0135124.
XX PR 21-MAY-1999; 99US-0135353.
XX PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
```

PR	63-AUG-1999;	99US-0147030;
PR	04-AUG-1999;	99US-0147204;
PR	04-AUG-1999;	99US-0147302;
PR	05-AUG-1999;	99US-0147192;
PR	05-AUG-1999;	99US-0147260;
PR	06-AUG-1999;	99US-0147303;
PR	06-AUG-1999;	99US-0147416;
PR	09-AUG-1999;	99US-0147493;
PR	09-AUG-1999;	99US-0147935;
PR	10-AUG-1999;	99US-0148171;
PR	11-AUG-1999;	99US-0148319;
PR	12-AUG-1999;	99US-0148341;
PR	13-AUG-1999;	99US-0148565;
PR	13-AUG-1999;	99US-0148684;
PR	16-AUG-1999;	99US-0149368;
PR	17-AUG-1999;	99US-0149175;
PR	18-AUG-1999;	99US-0149426;
PR	20-AUG-1999;	99US-0149722;
PR	20-AUG-1999;	99US-0149723;
PR	23-AUG-1999;	99US-0149929;
PR	23-AUG-1999;	99US-0149902;
PR	23-AUG-1999;	99US-0149930;
PR	25-AUG-1999;	99US-0150566;
PR	26-AUG-1999;	99US-0150884;
PR	27-AUG-1999;	99US-0151065;
PR	27-AUG-1999;	99US-0151066;
PR	30-AUG-1999;	99US-0151080;
PR	30-AUG-1999;	99US-0151303;
PR	31-SEP-1999;	99US-0151438;
PR	01-SEP-1999;	99US-0151930;
PR	07-SEP-1999;	99US-0152363;
PR	13-SEP-1999;	99US-0153070;
PR	13-SEP-1999;	99US-0153758;
PR	15-SEP-1999;	99US-0154018;
PR	16-SEP-1999;	99US-0154039;
PR	20-SEP-1999;	99US-0154779;
PR	22-SEP-1999;	99US-0155139;
PR	23-SEP-1999;	99US-0155346;
PR	24-SEP-1999;	99US-0155659;
PR	28-SEP-1999;	99US-0156458;
PR	29-SEP-1999;	99US-0156596;
PR	04-OCT-1999;	99US-0157117;
PR	05-OCT-1999;	99US-0157753;
PR	06-OCT-1999;	99US-0157865;
PR	07-OCT-1999;	99US-0158029;
PR	12-OCT-1999;	99US-0158232;
PR	12-OCT-1999;	99US-0158369;
PR	13-OCT-1999;	99US-0159293;
PR	13-OCT-1999;	99US-0159294;
PR	13-OCT-1999;	99US-0159295;
PR	14-OCT-1999;	99US-0159329;
PR	14-OCT-1999;	99US-0159330;
PR	14-OCT-1999;	99US-0159637;
PR	14-OCT-1999;	99US-0159638;
PR	18-OCT-1999;	99US-0159584;
PR	21-OCT-1999;	99US-0160741;
PR	21-OCT-1999;	99US-0160767;
PR	21-OCT-1999;	99US-0160768;
PR	21-OCT-1999;	99US-0160770;
PR	21-OCT-1999;	99US-0160814;
PR	22-OCT-1999;	99US-0160815;
PR	22-OCT-1999;	99US-0160980;
PR	22-OCT-1999;	99US-0160981;
PR	23-OCT-1999;	99US-0160989;
PR	25-OCT-1999;	99US-0161404;
PR	25-OCT-1999;	99US-0161405;
PR	25-OCT-1999;	99US-0161406;
PR	26-OCT-1999;	99US-0161359;
PR	26-OCT-1999;	99US-0161360;
PR	26-OCT-1999;	99US-0161361;
PR	28-OCT-1999;	99US-0161990;
PR	28-OCT-1999;	99US-0161992;

```

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 86.1%; Score 31; DB 21; Length 79;
Best Local Similarity 75.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
Db 20 TNNTLQTT 27

RESULT 9
AAG59706
ID AAG59706 standard; Protein; 90 AA.
XX AC AAG59706;
XX XX
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 77258.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX XX
XX PD 06-SEP-2000.
XX PF
XX PP
XX 25-FEB-2000; 2000EP-0301439.
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.

```

PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139460.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139462.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139463.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139750.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151065.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151066.
PR	22-JUN-1999;	99US-0139859.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140333.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140901.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142134.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	18-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160814.
PR	23-JUL-1999;	99US-0145218.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160980.
PR	26-JUL-1999;	99US-0145276.	PR	22-OCT-1999;	99US-0160981.
PR	27-JUL-1999;	99US-0145913.	PR	22-OCT-1999;	99US-0160989.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145951.	PR	25-OCT-1999;	99US-0161405.
PR	28-JUL-1999;	99US-0146386.	PR	25-OCT-1999;	99US-0161406.
PR	02-AUG-1999;	99US-0146388.	PR	26-OCT-1999;	99US-0161359.
PR	02-AUG-1999;	99US-0146389.	PR	26-OCT-1999;	99US-0161360.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161361.
PR	04-AUG-1999;	99US-0147204.	PR	28-OCT-1999;	99US-0161920.
PR	05-AUG-1999;	99US-0147302.	PR	28-OCT-1999;	99US-0161992.
PR	05-AUG-1999;	99US-0147192.	PR	28-OCT-1999;	99US-0161993.
PR	06-AUG-1999;	99US-0147260.	PR	29-OCT-1999;	99US-0162142.
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			

Query Match 86.1%; Score 31; DB 21; Length 90;  
Best Local Similarity 75.0%; Pred. No. 12;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8  
 DB 31 TNNFLQTT 38

RESULT 10  
 AAW98226  
 ID AAW98226 standard; Protein; 115 AA.  
 XX AC  
 XX AAW98226;  
 XX 31-MAR-1999 (first entry)  
 XX DT  
 XX H. pylori GHPO 1099 protein.  
 XX DE  
 XX GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;  
 KW peptic ulcer disease.  
 KW Helicobacter pylori.  
 XX OS  
 XX PN WO9843478-A1.  
 XX PD 08-OCT-1998.  
 XX PF 01-APR-1998; 98WO-US06371.  
 XX PR 29-JUL-1997; 97US-0902615.  
 XX PR 01-APR-1997; 97US-0833457.  
 XX PR 24-JUN-1997; 97US-0881227.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.  
 XX PI  
 XX Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;  
 XX DR WPI; 1998-542293/46.  
 XX DR N-PSDB; AAX13945.  
 XX PT New isolated Helicobacter polynucleotides - used to develop products  
 PT for the diagnosis, prevention and treatment of Helicobacter  
 PT infections and gastrointestinal diseases  
 XX PS Claim 8; Page 159-160; 2054pp; English.  
 XX CC This sequence represents a Helicobacter pylori GHPO protein of the  
 CC invention. The polypeptides can be used for preventing or treating  
 CC Helicobacter infections, and gastroduodenal diseases associated with  
 CC these infections, including acute, chronic, and atrophic gastritis, and  
 CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be  
 CC used for the production of antibodies. The products can also be used for  
 CC detection and diagnosis.  
 XX SQ Sequence 115 AA;  
 QY 1 TNNVLQ 6  
 DB 93 TNNVLQ 98

Query Match 83.3%; Score 30; DB 19; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
 DB 93 TNNVLQ 98

RESULT 11  
 AAY11062  
 ID AAY11062 standard; Protein; 116 AA.  
 XX AC  
 XX AAY11062;  
 XX 08-JUN-1999 (first entry)  
 XX DT  
 XX H. pylori ORF 06cp30603\_10744075\_c3\_136 secreted protein.  
 XX DE  
 XX PR

KW Vaccine; probe; diagnostic; ORF; cell envelope protein;  
 KW secreted protein; cytoplasmic protein; cellular protein.  
 XX Helicobacter pylori.  
 XX OS  
 XX PN WO9824475-A1.  
 XX PD 11-JUN-1998.  
 XX PF 05-DEC-1997; 97WO-US22104.  
 XX PR 14-JUL-1997; 97US-0891928.  
 XX PR 05-DEC-1996; 96US-0759625.  
 XX PR 25-MAR-1997; 97US-0823745.  
 XX (ASTR) ASTRA AB.  
 XX PI  
 XX Alm RA, Castriotta LM, Doig PC, Kabok Z, Smith D;  
 DR WPI; 1998-333051/29.  
 DR N-PSDB; AAX30591.  
 XX PT New isolated Helicobacter pylori nucleic acids - used to develop  
 PT products for the diagnosis, prevention and treatment of infection by  
 PT H. pylori and other Helicobacter species  
 XX PS Claims 37, 41; Page 224-225; 339pp; English.  
 XX CC Recombinant or substantially pure preparations of H. pylori polypeptides  
 CC are disclosed, together with the nucleic acids encoding them. In all,  
 CC 97 ORFs are shown. The proteins are variously cell envelope proteins,  
 CC cytoplasmic proteins, secreted proteins or other cellular proteins.  
 CC Vaccines containing the nucleic acids or proteins are claimed, as are  
 CC probes containing at least 8 nucleotides from the nucleic acid  
 CC sequences. The vaccines are useful for treating or reducing the risk of  
 CC H. pylori infections, and the probes can be used diagnostically for  
 CC detecting the presence of Helicobacter in a sample. The products are  
 CC also of use in screening for compounds having the ability to interfere  
 CC with the H. pylori life cycle or to inhibit H. pylori infection.  
 XX SQ Sequence 116 AA;  
 QY 1 TNNVLQ 6  
 DB 94 TNNVLQ 99

Query Match 83.3%; Score 30; DB 19; Length 116;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
 DB 94 TNNVLQ 99

RESULT 12  
 AAU72915  
 ID AAU72915 standard; Protein; 544 AA.  
 XX AC  
 XX AAU72915;  
 XX 12-MAR-2002 (first entry)  
 XX DT  
 XX Neisseria meningitidis virulence protein #5.  
 XX DE  
 XX Meningitis; virulence; gene; antibacterial; vaccine; veterinary;  
 KW infection; Gram-negative bacteria; antimicrobial.  
 KW Neisseria meningitidis.  
 XX OS  
 XX WO200185772-A2.  
 XX PN  
 XX 15-NOV-2001.  
 XX PD  
 XX 08-MAY-2001; 2001WO-GB02003.  
 XX PF  
 XX 08-MAY-2000; 2000GB-0011108.  
 XX PR

CC (MICR-) MICROSCIENCE LTD.  
CC Tang C;  
CC WPI; 2002-066593/09.  
DR N-PSDB; AAS97200.  
XX  
XX New peptide encoded by operon including virulence genes of Neisseria  
PT meningitidis, useful as vaccine component for treating or preventing  
PT meningitis and for identifying antimicrobial drug -  
XX Claim 4; Page 37-39; 423pp; English.  
PS  
XX The invention relates to a peptide (I) encoded by an operon (II) of  
CC Neisseria meningitidis including virulence genes, or a related molecule  
CC having a 40% sequence similarity at the peptide or nucleotide level in a  
CC Gram-negative bacterium, or its functional fragment, for therapeutic or  
CC diagnostic use. (I) and (II) are useful in the manufacture of a  
CC medicament for treating or preventing a condition (e.g., meningitis)  
CC associated with infection by Neisseria or Gram-negative bacteria. The  
CC product is useful for veterinary treatment and in a screening assay for  
CC the identification of an antimicrobial drug. The vaccines have  
CC prophylactic applications. AAU72911-AAU73014 represent N. meningitidis  
CC virulence proteins of the invention.  
XX  
SQ Sequence 544 AA;  
Query Match 83.3%; Score 30; DB 23; Length 544;  
Best Local Similarity 85.7%; Pred. No. 1.8e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 NNVLOXT 8  
DB 105 NNVLOTT 111  
RESULT 13  
ID ABB91881 standard; Protein; 960 AA.  
XX ABB91881;  
XX 31-MAY-2002 (first entry)  
XX Herbicidally active polypeptide SEQ ID NO 1092.  
XX Herbicidal; plant; agriculture; herbicide.  
XX Arabidopsis thaliana.  
XX WO200210210-A2.  
XX 07-FEB-2002.  
XX 28-AUG-2001; 2001WO-EF09892.  
XX 28-AUG-2001; 2001WO-EF09892.  
XX (FARB ) BAYER AG.  
XX Tietjen K, Weidler M;  
XX WPI; 2002-269010/31.  
XX Identifying plant target proteins for herbicidally active compounds,  
PT comprising aligning and comparing nucleic acid or amino acid sequences  
PT from plant with nucleic acid or amino acid sequences from non-plant  
PT organisms -  
XX Claim 5; SEQ ID NO 1092; 261pp + Sequence Listing; English.  
XX The invention relates to identifying target proteins  
CC

CC (ABB90790-ABB94016) for herbicidally active compounds, comprising  
CC aligning and comparing nucleic acid or amino acid sequences from plant  
CC with nucleic acid or amino acid sequences from non-plant organisms using  
CC suitable search parameters, where plant sequences having an E-value  
CC greater by a factor of 3 than the E-value of most similar non-plant  
CC sequences are selected. The polypeptides or nucleic acids encoding them  
CC are useful for identifying modulators. The identified modulators are  
CC useful as herbicides.  
XX Sequence 960 AA;  
SQ Query Match 83.3%; Score 30; DB 23; Length 960;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TNNVLQ 6  
DB 932 TNNVLQ 937  
RESULT 14  
ID AAG32057 standard; Protein; 524 AA.  
XX AAG32057;  
XX 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 38603.  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
XX 06-APR-1999; 99US-0128234.  
XX 18-APR-1999; 99US-0128714.  
XX 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
XX 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.  
XX 28-APR-1999; 99US-0130891.  
XX 30-APR-1999; 99US-0131449.  
XX 30-APR-1999; 99US-0132048.  
XX 30-APR-1999; 99US-0132407.  
XX 04-MAY-1999; 99US-0132484.  
XX 05-MAY-1999; 99US-0132485.  
XX 06-MAY-1999; 99US-0132486.  
XX 07-MAY-1999; 99US-0132487.  
XX 07-MAY-1999; 99US-0132863.  
XX 11-MAY-1999; 99US-0134256.  
XX 14-MAY-1999; 99US-0134218.  
XX 14-MAY-1999; 99US-0134219.  
XX 14-MAY-1999; 99US-0134221.  
XX 14-MAY-1999; 99US-0134370.  
XX 18-MAY-1999; 99US-0134768.  
XX 19-MAY-1999; 99US-0134941.  
XX 20-MAY-1999; 99US-0135124.  
XX 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136792.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159638.  
PR 21-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.

Wed Feb 26 15:45:40 2003

PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 80.6%; Score 29; DB 21; Length 524;  
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TNNVLQ 6  
 |||:  
 Db 44 TNNILQ 49

## RESULT 15

ABB91744  
 ID ABB91744 standard; Protein; 524 AA.

XX ABB91744;

XX 31-MAY-2002 (first entry)

XX Herbicidally active polypeptide SEQ ID NO 955.

XX Herbicidal; plant; agriculture; herbicide.

XX Arabidopsis thaliana.

XX WO200210210-A2.

XX 07-FEB-2002.

XX 28-AUG-2001; 2001WO-EP09892.

XX 28-AUG-2001; 2001WO-EP09892.

XX (FARB ) BAYER AG.

XX Tietjen K, Weidner M;

XX WPI; 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,  
 comprising aligning and comparing nucleic acid or amino acid sequences  
 from plant with nucleic acid or amino acid sequences from non-plant  
 organisms -

XX Claim 5; SEQ ID NO 955; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins  
 (ABB90790-ABB94016) for herbicidally active compounds, comprising  
 aligning and comparing nucleic acid or amino acid sequences from plant  
 with nucleic acid or amino acid sequences from non-plant organisms using  
 suitable search parameters, where plant sequences having an E-value  
 greater by a factor of 3 than the E-value of most similar non-plant  
 sequences are selected. The polypeptides or nucleic acids encoding them  
 are useful for identifying modulators. The identified modulators are  
 useful as herbicides.

XX Sequence 524 AA;

Query Match 80.6%; Score 29; DB 23; Length 524;  
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TNNVLQ 6  
 |||:  
 Db 44 TNNILQ 49

Search completed: February 26, 2003, 15:37:01  
 Job time : 36 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:38:16 ; Search time 32 Seconds  
(without alignments)  
9.430 Million cell updates/sec

Title: US-09-476-485A-24  
Perfect score: 36  
Sequence: 1 TNNVLQXT 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 174566 seqs, 37721826 residues

Total number of hits satisfying chosen parameters: 174566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications-AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	83.3	115	10	US-09-881-752A-52
2	27	75.0	311	10	US-09-898-416-10
3	27	75.0	731	9	US-09-738-626-4854
4	26	72.2	234	9	US-09-764-868-1025
5	26	72.2	500	10	US-09-323-998D-59
6	26	72.2	555	9	US-09-764-868-878
7	25	69.4	161	10	US-09-864-761-34751
8	25	69.4	166	10	US-09-760-541-6
9	25	69.4	258	9	US-09-738-626-4420
10	25	69.4	262	9	US-10-260-877-2
11	25	69.4	262	10	US-09-815-242-10954
12	25	69.4	308	10	US-09-925-299-930
13	25	69.4	308	10	US-09-765-272-122
14	25	69.4	522	10	US-09-965-313-7
15	25	69.4	546	9	US-09-905-291A-250
16	25	69.4	546	9	US-09-902-853-250
17	25	69.4	546	9	US-09-907-824-250
18	25	69.4	546	9	US-09-907-841-250
19	25	69.4	546	9	US-09-904-011-250
20	25	69.4	546	9	US-09-904-011-250
21	25	69.4	546	9	US-09-904-011-250
22	25	69.4	546	9	US-09-904-011-250
23	25	69.4	546	9	US-09-904-011-250
24	25	69.4	546	9	US-09-904-011-250
25	25	69.4	546	9	US-09-904-011-250
26	25	69.4	546	9	US-09-904-011-250
27	25	69.4	546	9	US-09-904-011-250
28	25	69.4	546	9	US-09-904-011-250
29	25	69.4	546	9	US-09-904-011-250
30	25	69.4	546	9	US-09-904-011-250
31	25	69.4	546	9	US-09-904-011-250
32	25	69.4	546	9	US-09-904-011-250
33	25	69.4	546	9	US-09-904-011-250
34	25	69.4	546	9	US-09-904-011-250
35	25	69.4	546	9	US-09-904-011-250
36	25	69.4	546	9	US-09-904-011-250
37	25	69.4	546	9	US-09-904-011-250
38	25	69.4	546	9	US-09-904-011-250
39	25	69.4	546	9	US-09-904-011-250
40	25	69.4	546	9	US-09-904-011-250
41	25	69.4	546	9	US-09-904-011-250
42	25	69.4	546	9	US-09-904-011-250
43	25	69.4	546	9	US-09-904-011-250
44	25	69.4	546	9	US-09-904-011-250
45	25	69.4	546	9	US-09-904-011-250

## ALIGNMENTS

## RESULT 1

US-09-881-752A-52  
; Sequence 52, Application US/09881752A  
; Patent No. US20020115078A1

## GENERAL INFORMATION:

; APPLICANT: Kleanthous, Harold  
; APPLICANT: Al-Garawi, Amal  
; APPLICANT: Miller, Charles  
; APPLICANT: Tomb, Jean-Francois  
; APPLICANT: Oomen, Raymond P.

; TITLE OF INVENTION: Identification of Polynucleotides

; TITLE OF INVENTION: Encoding No. US20020115078A1 Helicobacter Polypeptides in t

; FILE REFERENCE: 06132/041002

; CURRENT APPLICATION NUMBER: US/09/881,752A

; CURRENT FILING DATE: 2001-06-15

; PRIOR APPLICATION NUMBER: US 08/833,457

; PRIOR FILING DATE: 1997-04-01

; NUMBER OF SEQ ID NOS: 370

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 52

; LENGTH: 115

; TYPE: PRT

; ORGANISM: Helicobacter pylori

US-09-881-752A-52

Query Match 83.3%; Score 30; DB 10; Length 115;  
Best Local Similarity 100.0%; Pred. No. 6.9;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6

Db 93 TNNVLQ 98

## RESULT 2

US-09-898-416-10  
; Sequence 10, Application US/09898416

; Patent No. US20020076759A1

; GENERAL INFORMATION:

; APPLICANT: Dulac, Catherine

; APPLICANT: Axel, Richard

```
; TITLE OF INVENTION: Cloning Of Vertebrate Pheromone Receptors And Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0575/48557A/JPW/ADM
; CURRENT APPLICATION NUMBER: US/09/898,416
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 08/731,745
; PRIOR FILING DATE: 1996-10-18
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-898-416-10

Query Match          75.0%; Score 27; DB 10; Length 311;
Best Local Similarity 62.5%; Pred. No. 92;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TNNVLOXT 8
Db 163 TNNFMQVT 170

RESULT 3
US-09-738-626-4854
; Sequence 4854, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4854
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4854

Query Match          75.0%; Score 27; DB 9; Length 731;
Best Local Similarity 62.5%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TNNVLOXT 8
Db 639 TNNILAF 646

RESULT 4
US-09-764-868-1025
; Sequence 1025, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
```

```
; FILE REFERENCE: PTZ32
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1025
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (5)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (10)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-1025

Query Match          72.2%; Score 26; DB 9; Length 234;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNVLQ 6
Db 203 SNNVLQ 208

RESULT 5
US-09-323-998D-59
; Sequence 59, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 108172-05019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Nicotiana sp.
US-09-323-998D-59

Query Match          72.2%; Score 26; DB 10; Length 500;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NNVLQXT 8
Db 493 NNLLQDT 499

RESULT 6
US-09-764-868-878
; Sequence 878, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ32
```

; CURRENT APPLICATION NUMBER: US/09/764,868  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1510  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 878  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-868-878

Query Match 72.2%; Score 26; DB 9; Length 555;  
Best Local Similarity 83.3%; Pred. No. 2.8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
Db 524 SNNVLQ 529

RESULT 7  
US-09-864-761-34751  
; Sequence 34751, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecmica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 34751

LENGTH: 81  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
; OTHER INFORMATION: MAP TO AP000352.2  
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6.3  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 9.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.4  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.9  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 6.7  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.3  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.2  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.5  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 9.9  
; OTHER INFORMATION: SWISSPROT HIT: Q9Y6J0, EVALUE 2.00e-39  
; OTHER INFORMATION: EST\_HUMAN HIT: A0120832.1, EVALUE 4.80e+00  
US-09-864-761-34751

Query Match 69.4%; Score 25; DB 10; Length 81;  
Best Local Similarity 66.7%; Pred. No. 59;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
Db 76 TNNLIQ 81

RESULT 8  
US-09-760-541-6  
; Sequence 6, Application US/09760541  
; Patent No. US20010019713A1  
; GENERAL INFORMATION:  
; APPLICANT: Blazar, Beverly A.  
; APPLICANT: Webb, Andrew C.  
; TITLE OF INVENTION: NO. US20010019713A1 Interleukin Compositions and Methods  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David R. Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/760,541  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/287,387  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saliwanchik, David R.  
; REGISTRATION NUMBER: 31,794  
; REFERENCE/DOCKET NUMBER: BLAZ-1.C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 904-375-8100  
; TELEFAX: 904-372-5800  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 166 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-760-541-6

```
Query Match          69.4%; Score 25; DB 10; Length 166;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 NNVLQXT 8
Db 55 NNLSST 61

RESULT 9
US-09-738-626-4420
; Sequence 4420, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4420
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4420

Query Match          69.4%; Score 25; DB 9; Length 258;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVL 5
Db 223 TNNVL 227

RESULT 10
US-10-260-877-2
; Sequence 2, Application US/10260877
; Publication No. US20030021813A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Chovan, Linda E.
; APPLICANT: Hessler, Paul E.
; APPLICANT: Reich, Karl A.
; TITLE OF INVENTION: ESSENTIAL BACTERIA GENES AND GENOME
; TITLE OF INVENTION: SCANNING IN HAEMOPHILUS INFLUENZAE FOR THE IDENTIFICATION OF
; FILE REFERENCE: 6565.US.P1
; CURRENT APPLICATION NUMBER: US/10/260,877
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US/09/649,145
; PRIOR FILING DATE: 2000-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 262
; TYPE: PRT
```

```
; ORGANISM: H. influenzae
US-10-260-877-2

Query Match          69.4%; Score 25; DB 9; Length 262;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVL 5
Db 59 TNNVL 63

RESULT 11
US-09-815-242-10954
; Sequence 10954, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10954
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-815-242-10954

Query Match          69.4%; Score 25; DB 10; Length 262;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVL 5
Db 59 TNNVL 63

RESULT 12
US-09-925-299-930
; Sequence 930, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
```

```
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 930
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (110)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (152)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (225)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-299-930

Query Match          69.4%; Score 25; DB 10; Length 308;
Best Local Similarity 50.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
|:|:|
Db 241 TDNLQAT 248

RESULT 13
US-09-765-272-122
; Sequence 122, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 520 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

```
; SEQUENCE DESCRIPTION: SEQ ID NO: 122:
US-09-765-272-122

Query Match          69.4%; Score 25; DB 10; Length 520;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVL 5
|:|:|
Db 507 TNNVL 511

RESULT 14
US-09-965-313-7
; Sequence 7, Application US/09965313
; Patent No. US20020090680A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: No. US20020090680A1 IL-9/TL-2 Receptor-Like Molecules
; FILE OF INVENTION: and Uses Thereof
; FILE REFERENCE: 5800-17A
; CURRENT APPLICATION NUMBER: US/09/965,313
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 09/313,913
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-965-313-7

Query Match          69.4%; Score 25; DB 10; Length 522;
Best Local Similarity 66.7%; Pred. No. 4.4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
|:|:|
Db 58 TNNILR 63

RESULT 15
US-09-905-291A-250
; Sequence 250, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
```

; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,291A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 250
; LENGTH: 546
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-905-291A-250

Query Match 69.4%; Score 25; DB 9; Length 546;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNVLQ 6
Db 476 NNVLQ 480

Search completed: February 26, 2003, 15:42:46
Job time : 33 secs

GenCore version 5.1.3

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:37:06 ; Search time 139 Seconds

(without alignments)

37.107 Million cell updates/sec

Title: US-09-476-485A-24

Perfect score: 36

Sequence: 1 TNNVLQXT 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pcp.\*  
 2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pcp.\*  
 3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pcp.\*  
 4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pcp.\*  
 5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pcp.\*  
 6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pcp.\*  
 7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pcp.\*  
 8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pcp.\*  
 9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pcp.\*  
 10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pcp.\*  
 11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pcp.\*  
 12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pcp.\*  
 13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pcp.\*  
 14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pcp.\*  
 15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pcp.\*  
 16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pcp.\*  
 17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pcp.\*  
 18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pcp.\*  
 19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pcp.\*  
 20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pcp.\*  
 21: /cgn2\_6/ptodata/1/paa/US097\_COMB.pcp.\*  
 22: /cgn2\_6/ptodata/1/paa/US098\_COMB.pcp.\*  
 23: /cgn2\_6/ptodata/1/paa/US099\_COMB.pcp.\*  
 24: /cgn2\_6/ptodata/1/paa/US100\_COMB.pcp.\*  
 25: /cgn2\_6/ptodata/1/paa/US101\_COMB.pcp.\*  
 26: /cgn2\_6/ptodata/1/paa/US102\_COMB.pcp.\*  
 27: /cgn2\_6/ptodata/1/paa/US60\_COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	94.4	8	1	PCT-US98-13046-24
2	34	94.4	8	18	US-09-476-485A-24
3	34	94.4	8	24	US-10-045-353-24
4	34	94.4	123	18	US-09-476-485A-50
5	34	94.4	264	18	US-09-476-485A-2
6	34	94.4	270	1	PCT-US98-13046-2

7	34	94.4	270	24	US-10-045-353-2	Sequence 2, Appli
8	34	94.4	272	21	US-09-791-537-1173	Sequence 1173, Ap
9	34	94.4	279	18	US-09-476-485A-56	Sequence 56, Appl
10	34	94.4	279	21	US-09-791-537-109055	Sequence 109055,
11	34	94.4	286	1	PCT-US98-13046-23	Sequence 23, Appl
12	34	94.4	286	18	US-09-476-485A-23	Sequence 23, Appl
13	34	94.4	286	24	US-10-045-353-23	Sequence 6, Appli
14	34	94.4	303	18	US-09-476-485A-6	Sequence 77260, A
15	31	86.1	73	19	US-09-513-996A-77259	Sequence 1503, Ap
16	31	86.1	74	19	US-09-513-996A-77258	Sequence 1502, Ap
17	31	86.1	79	19	US-09-513-996A-77258	Sequence 1502, Ap
18	31	86.1	80	19	US-09-513-996A-77258	Sequence 1502, Ap
19	31	86.1	90	19	US-09-513-996A-77258	Sequence 1502, Ap
20	31	86.1	91	19	US-09-513-996A-77258	Sequence 1502, Ap
21	31	86.1	104	12	US-08-827-356-4896	Sequence 6497, Ap
22	31	86.1	104	20	US-09-611-529-6497	Sequence 79655, A
23	30	83.3	25	21	US-09-724-429-79655	Sequence 79656, A
24	30	83.3	25	21	US-09-724-429-79655	Sequence 80231, A
25	30	83.3	25	21	US-09-724-429-80231	Sequence 80232, A
26	30	83.3	25	21	US-09-724-429-80232	Sequence 81202, A
27	30	83.3	25	21	US-09-724-429-81202	Sequence 81203, A
28	30	83.3	25	21	US-09-724-429-81203	Sequence 81204, A
29	30	83.3	25	21	US-09-724-429-81204	Sequence 82066, A
30	30	83.3	25	21	US-09-724-429-82066	Sequence 82067, A
31	30	83.3	25	21	US-09-724-429-82067	Sequence 82068, A
32	30	83.3	25	21	US-09-724-429-82068	Sequence 83409, A
33	30	83.3	25	21	US-09-724-429-83409	Sequence 83410, A
34	30	83.3	25	21	US-09-724-429-83410	Sequence 83411, A
35	30	83.3	25	21	US-09-724-429-83411	Sequence 83412, A
36	30	83.3	25	21	US-09-724-429-83412	Sequence 84561, A
37	30	83.3	25	21	US-09-724-429-84561	Sequence 84562, A
38	30	83.3	25	21	US-09-724-429-84562	Sequence 84563, A
39	30	83.3	25	21	US-09-724-429-84563	Sequence 84564, A
40	30	83.3	25	21	US-09-724-429-84564	Sequence 84565, A
41	30	83.3	25	21	US-09-724-429-84565	Sequence 84566, A
42	30	83.3	25	21	US-09-724-429-84566	Sequence 84567, A
43	30	83.3	25	21	US-09-724-429-84567	Sequence 84568, A
44	30	83.3	25	21	US-09-724-429-84568	Sequence 84569, A
45	30	83.3	25	21	US-09-724-429-84569	Sequence 84570, A

## ALIGNMENTS

RESULT 1

PCT-US98-13046-24

; Sequence 24, Application PC/TUS9813046

; GENERAL INFORMATION:

; APPLICANT: Colucci et al.

; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hoffmann &amp; Baron, LLP

; STREET: 6900 Jericho Turnpike

; CITY: Syosset

; STATE: New York

; COUNTRY: USA

; ZIP: 11791

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage

; COMPUTER: IBM compatible

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WordPerfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US98/13046

; FILING DATE: June 23, 1998

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/881,189

; FILING DATE: June 24, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Feit, Irving N.

; REGISTRATION NUMBER: 28,601

```

; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; PCT-US98-13046-24

Query Match          94.4%; Score 34; DB 1; Length 8;
Best Local Similarity 87.5%; Pred. No. 4.2e+06;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
Db 1 TNNVLQVT 8

RESULT 2
US-09-476-485A-24
; Sequence 24, Application US/09476485A
; GENERAL INFORMATION:
; APPLICANT: Colucci, M. Gabriella
; APPLICANT: Chrispeels, Maarten J.
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for
; and Products of Their Use
; FILE REFERENCE: 108236.119
; CURRENT APPLICATION NUMBER: US/09/476.485A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/881,189
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Dolichos lablab
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (7)..(7)
; OTHER INFORMATION: Amino acid 7 is Xaa wherein Xaa = any amino acid.
US-09-476-485A-24

Query Match          94.4%; Score 34; DB 18; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
Db 1 TNNVLQXT 8

RESULT 3
US-10-045-353-24
; Sequence 24, Application US/10045353
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; PROGENITOR CELL PRESERVATION FACTOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible

; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/045.353
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/881,189
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Felt, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-10-045-353-24

Query Match          94.4%; Score 34; DB 24; Length 8;
Best Local Similarity 87.5%; Pred. No. 4.2e+06;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
Db 1 TNNVLQVT 8

RESULT 4
US-09-476-485A-50
; Sequence 50, Application US/09476485A
; GENERAL INFORMATION:
; APPLICANT: Colucci, M. Gabriella
; APPLICANT: Chrispeels, Maarten J.
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for
; and Products of Their Use
; FILE REFERENCE: 108236.119
; CURRENT APPLICATION NUMBER: US/09/476.485A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/881,189
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Beta-subunit of DI-FRIL.
US-09-476-485A-50

Query Match          94.4%; Score 34; DB 18; Length 123;
Best Local Similarity 87.5%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
Db 27 TNNVLQVT 34

RESULT 5
US-09-476-485A-2
; Sequence 2, Application US/09476485A
; GENERAL INFORMATION:
; APPLICANT: Colucci, M. Gabriella
; APPLICANT: Chrispeels, Maarten J.
; APPLICANT: Moore, Jeffrey G.

```



; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for  
 ; TITLE OF INVENTION: and Products of Their Use  
 ; FILE REFERENCE: 108236.119  
 ; CURRENT APPLICATION NUMBER: US/09/476.485A  
 ; CURRENT FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: US 08/881,189  
 ; PRIOR FILING DATE: 1997-06-24  
 ; NUMBER OF SEQ ID NOS: 57  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 264  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: D1-FRIL.  
 US-09-476-485A-2

Query Match 94.4%; Score 34; DB 18; Length 264;  
 Best Local Similarity 87.5%; Pred. No. 42;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8  
 | | | | | | |

Db 27 TNNVLQXT 34

RESULT 6  
 PCT-US98-13046-2  
 ; Sequence 2, Application PC/TUS9813046  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Colucci et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED  
 ; NUMBER OF SEQUENCES: 24  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hoffmann & Baron, LLP  
 ; STREET: 6900 Jericho Turnpike  
 ; CITY: Syosset  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 11791  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
 ; COMPUTER: IBM compatible  
 ; OPERATING SYSTEM: MS-DOS  
 ; SOFTWARE: WordPerfect  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US98/13046  
 ; FILING DATE: June 23, 1998  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/881,189  
 ; FILING DATE: June 24, 1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Feit, Irving N.  
 ; REGISTRATION NUMBER: 28,601  
 ; REFERENCE/DOCKET NUMBER: 381-44 PCT  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (516) 822-3550  
 ; TELEFAX: (516) 822-3582  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 270 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide

PCT-US98-13046-2  
 Query Match 94.4%; Score 34; DB 1; Length 270;  
 Best Local Similarity 87.5%; Pred. No. 43;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8  
 | | | | | | |

Db 27 TNNVLQXT 34  
 RESULT 7  
 US-10-045-353-2  
 ; Sequence 2, Application US/10045353  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Colucci et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED  
 ; PROGENITOR CELL PRESERVATION FACTOR  
 ; NUMBER OF SEQUENCES: 24  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Hoffmann & Baron, LLP  
 ; STREET: 350 Jericho Turnpike  
 ; CITY: Jericho  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 11753  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
 ; COMPUTER: IBM compatible  
 ; OPERATING SYSTEM: MS-DOS  
 ; SOFTWARE: WordPerfect  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/10/045.353  
 ; FILING DATE: 29-Oct-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/881,189  
 ; FILING DATE: <Unknown>  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Feit, Irving N.  
 ; REGISTRATION NUMBER: 28,601  
 ; REFERENCE/DOCKET NUMBER: 381-44 PCT  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (516) 822-3550  
 ; TELEFAX: (516) 822-3582  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 270 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 US-10-045-353-2

Query Match 94.4%; Score 34; DB 24; Length 270;  
 Best Local Similarity 87.5%; Pred. No. 43;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8  
 | | | | | | |

Db 27 TNNVLQXT 34

RESULT 8  
 US-09-791-537-1173  
 ; Sequence 1173, Application US/09791537  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Blonmox, Inc.  
 ; APPLICANT: Debe, Derek  
 ; APPLICANT: Danzer, Joseph  
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
 ; FILE REFERENCE: 261/210  
 ; CURRENT APPLICATION NUMBER: US/09/791,537  
 ; CURRENT FILING DATE: 2001-02-22  
 ; NUMBER OF SEQ ID NOS: 153055  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 1173  
 ; LENGTH: 272  
 ; TYPE: PRT  
 ; ORGANISM: Dolichos lablab

## US-09-791-537-1173

Query Match 94.4%; Score 34; DB 21; Length 272;  
Best Local Similarity 87.5%; Pred. No. 43;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8  
| | | | |  
Db 35 TNNVLQVT 42

## RESULT 9

US-09-476-485A-56  
; Sequence 56, Application US/09476485A  
; GENERAL INFORMATION:  
; APPLICANT: Colucci, M. Gabriella  
; APPLICANT: Chrispeels, Maarten J.  
; APPLICANT: Moore, Jeffrey G.  
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for  
; FILE REFERENCE: 108236.119  
; CURRENT APPLICATION NUMBER: US/09/476.485A  
; CURRENT FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 08/881.189  
; PRIOR FILING DATE: 1997-06-24  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 56  
; LENGTH: 279  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: PVPRIIL.  
US-09-476-485A-56

Query Match 94.4%; Score 34; DB 18; Length 279;  
Best Local Similarity 87.5%; Pred. No. 44;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8  
| | | | |  
Db 27 TNNVLQLT 34

## RESULT 10

US-09-791-537-109055  
; Sequence 109055, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 109055  
; LENGTH: 279  
; TYPE: PRT  
; ORGANISM: Phaseolus vulgaris  
US-09-791-537-109055

Query Match 94.4%; Score 34; DB 21; Length 279;  
Best Local Similarity 87.5%; Pred. No. 44;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8  
| | | | |  
Db 27 TNNVLQLT 34

## RESULT 11

PCT-US98-13046-23  
; Sequence 23, Application PC/TUS9813046  
; GENERAL INFORMATION:  
; APPLICANT: Colucci et al.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron, LLP  
; STREET: 6900 Jericho Turnpike  
; CITY: Syosset  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11791  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US98/13046  
; FILING DATE: June 23, 1998  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/881,189  
; FILING DATE: June 24, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: 381-44 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 286 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
PCT-US98-13046-23

Query Match 94.4%; Score 34; DB 1; Length 286;  
Best Local Similarity 87.5%; Pred. No. 46;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8  
| | | | |  
Db 49 TNNVLQVT 56

## RESULT 12

US-09-476-485A-23  
; Sequence 23, Application US/09476485A  
; GENERAL INFORMATION:  
; APPLICANT: Colucci, M. Gabriella  
; APPLICANT: Chrispeels, Maarten J.  
; APPLICANT: Moore, Jeffrey G.  
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for  
; FILE REFERENCE: 108236.119  
; CURRENT APPLICATION NUMBER: US/09/476.485A  
; CURRENT FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 08/881,189  
; PRIOR FILING DATE: 1997-06-24  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 23  
; LENGTH: 286  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: SpDLA.  
US-09-476-485A-23

Query Match 94.4%; Score 34; DB 18; Length 286;  
Best Local Similarity 87.5%; Pred. No. 46;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8  
| | | | |  
Db 49 TNNVLOXT 56

RESULT 13  
US-10-045-353-23  
; Sequence 23, Application US/10045353  
; GENERAL INFORMATION:  
; APPLICANT: Colucci et al.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED  
; PROGENITOR CELL PRESERVATION FACTOR  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron, LLP  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11753

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/045,353  
FILING DATE: 29-Oct-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/881,189  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: 381-44 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 822-3550  
TELEFAX: (516) 822-3582  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 286 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 23:  
US-10-045-353-23

Query Match 94.4%; Score 34; DB 24; Length 286;  
Best Local Similarity 87.5%; Pred. No. 46;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8  
| | | | |  
Db 49 TNNVLOXT 56

RESULT 14  
US-09-476-485A-6  
; Sequence 6, Application US/09476485A  
; GENERAL INFORMATION:  
; APPLICANT: Colucci, M. Gabriella  
; APPLICANT: Chrispeels, Maarten J.  
; APPLICANT: Moore, Jeffrey G.  
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for  
; FILE OF INVENTION: and Products of Their Use  
; FILE REFERENCE: 108236.119  
; CURRENT APPLICATION NUMBER: US/09/476,485A  
; CURRENT FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 08/881,189

; PRIOR FILING DATE: 1997-06-24  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 6  
; LENGTH: 303  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Pv-FRIL.  
US-09-476-485A-6

Query Match 94.4%; Score 34; DB 18; Length 303;  
Best Local Similarity 87.5%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8  
| | | | |  
Db 27 TNNVLOXT 34

RESULT 15  
US-09-513-996A-77260  
; Sequence 77260, Application US/09513996A  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE OF INVENTION: ENCODED THEREBY  
; FILE REFERENCE: 2750-709P  
; CURRENT APPLICATION NUMBER: US/09/513,996A  
; CURRENT FILING DATE: 2000-02-25  
; NUMBER OF SEQ ID NOS: 81028  
; SEQ ID NO 77260  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 1..73 /  
; OTHER INFORMATION: any n or Xaa = unknown  
; FEATURE:  
; OTHER INFORMATION: Location 1..73 / Ceres Seq. ID 2104968  
US-09-513-996A-77260

Query Match 86.1%; Score 31; DB 19; Length 73;  
Best Local Similarity 75.0%; Pred. No. 47;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8  
| | | | |  
Db 14 TNNVLOXT 21

Search completed: February 26, 2003, 15:41:37  
Job time : 140 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:37:41 ; Search time 23 Seconds  
(without alignments)  
31.790 Million cell updates/sec

Title: US-09-476-485A-24  
Perfect score: 36  
Sequence: 1 TNNVLQXT 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 486122 seqs, 91396495 residues

Total number of hits satisfying chosen parameters: 486122

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*

- 1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	34	94.4	8	US-10-190-258A-11	Sequence 11, Appl
2	34	94.4	264	US-10-083-936B-2	Sequence 2, Appl
3	34	94.4	264	US-10-190-258A-2	Sequence 2, Appl
4	34	94.4	279	US-10-190-258A-6	Sequence 6, Appl
5	34	94.4	303	US-10-083-936B-6	Sequence 6, Appl
6	31	86.1	104	US-09-950-084-6497	Sequence 6497, Ap
7	27	75.0	328	US-10-287-274-376	Sequence 376, App
8	27	75.0	365	PCT-US02-35143-13	Sequence 13, Appl
9	27	75.0	365	US-10-286-326-13	Sequence 13, Appl
10	27	75.0	1912	PCT-US02-21361-48	Sequence 48, Appl
11	27	75.0	1912	US-10-188-186-48	Sequence 48, Appl
12	26	72.2	257	PCT-US02-36123-1120	Sequence 1120, Ap
13	26	72.2	262	US-09-950-084-4917	Sequence 4917, Ap
14	26	72.2	310	US-10-157-104-77	Sequence 77, Appl
15	26	72.2	310	US-10-005-041A-49	Sequence 49, Appl
16	26	72.2	310	US-10-005-041B-49	Sequence 49, Appl
17	26	72.2	313	US-10-157-104-73	Sequence 73, Appl
18	26	72.2	313	US-10-005-041A-50	Sequence 50, Appl
19	26	72.2	313	US-10-005-041A-108	Sequence 108, App
20	26	72.2	313	US-10-005-041B-50	Sequence 50, Appl
21	26	72.2	313	US-10-005-041B-108	Sequence 108, App
22	26	72.2	357	PCT-US02-36123-1122	Sequence 1122, Ap
23	26	72.2	357	US-10-157-104-75	Sequence 75, Appl
24	26	72.2	357	US-10-157-104-76	Sequence 76, Appl
25	26	72.2	357	US-10-005-041A-12	Sequence 12, Appl
26	26	72.2	357	US-10-005-041A-47	Sequence 47, Appl

27 26 72.2 357 6 US-10-005-041A-48 Sequence 48, Appl  
28 26 72.2 357 6 US-10-005-041A-109 Sequence 109, App  
29 26 72.2 357 6 US-10-005-041A-111 Sequence 111, App  
30 26 72.2 357 6 US-10-005-041B-12 Sequence 12, Appl  
31 26 72.2 357 6 US-10-005-041B-47 Sequence 47, Appl  
32 26 72.2 357 6 US-10-005-041B-48 Sequence 48, Appl  
33 26 72.2 357 6 US-10-005-041B-109 Sequence 109, App  
34 26 72.2 357 6 US-10-005-041B-111 Sequence 111, App  
35 26 72.2 400 1 PCT-US02-36123-1124 Sequence 22, Appl  
36 26 72.2 400 1 PCT-US02-36123-1124 Sequence 1124, Ap  
37 26 72.2 418 1 PCT-US02-00667A-53 Sequence 53, Appl  
38 26 72.2 631 6 US-10-310-154-436 Sequence 436, App  
39 26 72.2 779 6 US-10-072-012-564 Sequence 564, App  
40 26 72.2 779 6 US-10-072-012-731 Sequence 731, App  
41 26 72.2 1136 6 US-10-072-012-280 Sequence 280, App  
42 26 72.2 1136 6 US-10-072-012-729 Sequence 729, App  
43 26 72.2 1165 6 US-10-072-012-730 Sequence 730, App  
44 26 72.2 1178 6 US-10-310-154-551 Sequence 551, App  
45 25 69.4 64 6 US-10-320-646-31 Sequence 31, Appl

## ALIGNMENTS

RESULT 1  
US-10-190-258A-11  
; Sequence 11, Application US/10190258A  
; GENERAL INFORMATION:  
; APPLICANT: Moore, Jeffrey G  
; TITLE OF INVENTION: Dendritic Cell Isolation Methods  
; FILE REFERENCE: 108236.132  
; CURRENT APPLICATION NUMBER: US/10/190, 258A  
; CURRENT FILING DATE: 2002-07-03  
; PRIOR APPLICATION NUMBER: US 60/303,265  
; PRIOR FILING DATE: 2001-07-05  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Dolichos lablab  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 7  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-10-190-258A-11

Query Match 94.4%; Score 34; DB 6; Length 8;  
Best Local Similarity 100.0%; Pred. No. 4e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8  
Db 1 TNNVLQXT 8

RESULT 2  
US-10-083-936B-2  
; Sequence 2, Application US/10083936B  
; GENERAL INFORMATION:  
; APPLICANT: Moore, Jeffrey G.  
; TITLE OF INVENTION: Compositions and Methods for Protecting Tissues and  
; FILE REFERENCE: 108236.130  
; CURRENT APPLICATION NUMBER: US/10/083,936B  
; CURRENT FILING DATE: 2002-10-22  
; PRIOR APPLICATION NUMBER: US 60/271,666  
; PRIOR FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: US 60/302,716  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2

```

; LENGTH: 264
; TYPE: PRT
; ORGANISM: Dolichos lablab
US-10-083-936B-2

Query Match          94.4%; Score 34; DB 6; Length 264;
Best Local Similarity 87.5%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
Db 27 TNNVLQVT 34

RESULT 3
US-10-190-258A-2
; Sequence 2, Application US/10190258A
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G
; TITLE OF INVENTION: Dendritic Cell Isolation Methods
; FILE REFERENCE: 108236.132
; CURRENT APPLICATION NUMBER: US/10/190,258A
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/303,265
; PRIOR FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Dolichos lablab
US-10-190-258A-2

Query Match          94.4%; Score 34; DB 6; Length 264;
Best Local Similarity 87.5%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
Db 27 TNNVLQVT 34

RESULT 4
US-10-190-258A-6
; Sequence 6, Application US/10190258A
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G
; TITLE OF INVENTION: Dendritic Cell Isolation Methods
; FILE REFERENCE: 108236.132
; CURRENT APPLICATION NUMBER: US/10/190,258A
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/303,265
; PRIOR FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Phaseolus vulgaris
US-10-190-258A-6

Query Match          94.4%; Score 34; DB 6; Length 279;
Best Local Similarity 87.5%; Pred. No. 1.6;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
Db 27 TNNVLQVT 34

RESULT 5
US-10-083-936B-6
; Sequence 6, Application US/10083936B
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G
; TITLE OF INVENTION: Compositions and Methods for Protecting Tissues and
; FILE REFERENCE: 108236.130
; CURRENT APPLICATION NUMBER: US/10/083,936B
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: US 60/271,666
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: US 60/302,716
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Phaseolus vulgaris
US-10-083-936B-6

Query Match          94.4%; Score 34; DB 6; Length 303;
Best Local Similarity 87.5%; Pred. No. 1.8;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
Db 27 TNNVLQVT 34

RESULT 6
US-09-950-084-6497
; Sequence 6497, Application US/09950084
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Roberta S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
; FILE REFERENCE: 1034/1C963US2
; CURRENT APPLICATION NUMBER: US/09/950,084
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 09/417,811
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: US 09/353,718
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/266,557
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,555
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,541
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/037,934
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 09/036,720
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,338
; PRIOR FILING DATE: 1998-03-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO 6497
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-950-084-6497

Query Match          86.1%; Score 31; DB 5; Length 104;
Best Local Similarity 75.0%; Pred. No. 2.6;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8

```

Db 71 TNNVLEIT 78

## RESULT 7

US-10-287-274-376  
; Sequence 376, Application US/10287274  
; GENERAL INFORMATION:  
; APPLICANT: Forsyth, R. Allyn  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERAPY  
; FILE REFERENCE: ELITRA.008DVI  
; CURRENT APPLICATION NUMBER: US/10/287,274  
; CURRENT FILING DATE: 2002-10-31  
; PRIOR APPLICATION NUMBER: US 60/164415  
; PRIOR FILING DATE: 1999-11-09  
; PRIOR APPLICATION NUMBER: US 09/711164  
; PRIOR FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 469  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 376  
; LENGTH: 328  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-10-287-274-376

Query Match 75.0%; Score 27; DB 6; Length 328;

Best Local Similarity 66.7%; Pred. No. 94;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6

Db 83 TNNIIQ 88

## RESULT 8

PCT-US02-35143-13  
; Sequence 13, Application PC/TUS0235143  
; GENERAL INFORMATION:  
; APPLICANT: San, Ka-Yui  
; APPLICANT: Berrios-Rivera, Susana  
; APPLICANT: Bennett, George  
; TITLE OF INVENTION: Recycling System for Manipulation of Intracellular NADH Availability  
; FILE REFERENCE: P02328WO  
; CURRENT APPLICATION NUMBER: PCT/US02/35143  
; CURRENT FILING DATE: 2002-11-01  
; PRIOR APPLICATION NUMBER: US 60/335,371  
; PRIOR FILING DATE: 2001-11-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 13  
; LENGTH: 365  
; TYPE: PRT  
; ORGANISM: Rhodococcus  
PCT-US02-35143-13

Query Match 75.0%; Score 27; DB 1; Length 365;

Best Local Similarity 83.3%; Pred. No. 1.1e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6

Db 338 TNNLLQ 343

## RESULT 9

US-10-286-326-13  
; Sequence 13, Application US/10286326  
; GENERAL INFORMATION:  
; APPLICANT: San, Ka-Yui  
; APPLICANT: Berrios-Rivera, Susana  
; APPLICANT: Bennett, George

; TITLE OF INVENTION: Recycling System for Manipulation of Intracellular NADH Availability  
; FILE REFERENCE: P02328US1  
; CURRENT APPLICATION NUMBER: US/10/286,326  
; CURRENT FILING DATE: 2002-11-01  
; PRIOR APPLICATION NUMBER: US 60/335,371  
; PRIOR FILING DATE: 2001-11-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 13  
; LENGTH: 365  
; TYPE: PRT  
; ORGANISM: Rhodococcus  
US-10-286-326-13

Query Match 75.0%; Score 27; DB 6; Length 365;

Best Local Similarity 83.3%; Pred. No. 1.1e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6

Db 338 TNNLLQ 343

## RESULT 10

PCT-US02-21361-48  
; Sequence 48, Application PC/TUS0221361  
; GENERAL INFORMATION:  
; APPLICANT: CuraGen Corporation, et al.  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-397C  
; CURRENT APPLICATION NUMBER: PCT/US02/21361  
; CURRENT FILING DATE: 2002-07-03  
; PRIOR APPLICATION NUMBER: 60/303046  
; PRIOR FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 60/360814  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/303828  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/323380  
; PRIOR FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 60/361133  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/304016  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 60/304502  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 60/305262  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 60/373881  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/305673  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: Custom

Query Match 75.0%; Score 27; DB 1; Length 1912;

Best Local Similarity 71.4%; Pred. No. 7.5e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNVLQXT 8

Db 1355 NNVIQRT 1361

## RESULT 11

US-10-188-186-48  
; Sequence 48, Application US/10188186

US-09-950-084-4917  
; GENERAL INFORMATION:  
; APPLICANT: Anderson et al.  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-397C  
; CURRENT APPLICATION NUMBER: US/10/188,186  
; CURRENT FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: 60/303046  
; PRIOR FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 60/360814  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/303828  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/323380  
; PRIOR FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 60/361133  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/304016  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 60/304502  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 60/305262  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 60/373881  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/305673  
; PRIOR FILING DATE: 2001-07-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: Custom  
; SEQ ID NO 48  
; LENGTH: 1912  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-188-186-48

Query Match 75.0%; Score 27; DB 6; Length 1912;  
Best Local Similarity 71.4%; Pred. No. 7.5e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NNVLOXT 8  
Db 1355 NNVLOXT 1361  
RESULT 12  
PCT-US02-36123-1120  
; GENERAL INFORMATION:  
; APPLICANT: American Cyanamid Company, and Fletcher, Leah D., McMichael, John C.,  
; TITLE OF INVENTION: Allocoococcus otitis Open Reading Frames (ORFs) Encoding Polypeptides  
; FILE REFERENCE: Antigen, Immunogenic Compositions and Uses Thereof  
; CURRENT APPLICATION NUMBER: PCT/US02/36123  
; CURRENT FILING DATE: 2003-01-02  
; NUMBER OF SEQ ID NOS: 6653  
; SOFTWARE: Patent version 3.1  
; SEQ ID NO 1120  
; LENGTH: 257  
; TYPE: PRT  
; ORGANISM: Allocoococcus otitis  
PCT-US02-36123-1120

Query Match 72.2%; Score 26; DB 1; Length 257;  
Best Local Similarity 71.4%; Pred. No. 1.2e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NNVLOXT 8  
Db 166 NNVLOXT 172  
RESULT 13

US-09-950-084-4917  
; GENERAL INFORMATION:  
; APPLICANT: George H. Shimer, Jr.  
; APPLICANT: George H. Miller  
; APPLICANT: Roberta S. Hare  
; APPLICANT: Karen J. Shaw  
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods  
; FILE REFERENCE: 1034/1C963US2  
; CURRENT APPLICATION NUMBER: US/09/950,084  
; CURRENT FILING DATE: 2001-09-10  
; PRIOR APPLICATION NUMBER: US 09/417,811  
; PRIOR FILING DATE: 1999-10-14  
; PRIOR APPLICATION NUMBER: US 09/353,718  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: US 09/266,557  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 09/266,556  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 09/266,555  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 09/266,542  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 09/266,541  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 09/037,934  
; PRIOR FILING DATE: 1998-03-10  
; PRIOR APPLICATION NUMBER: US 09/036,720  
; PRIOR FILING DATE: 1998-03-06  
; PRIOR APPLICATION NUMBER: US 09/036,338  
; PRIOR FILING DATE: 1998-03-06  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 7451  
; SEQ ID NO 4917  
; LENGTH: 262  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-09-950-084-4917

Query Match 72.2%; Score 26; DB 5; Length 262;  
Best Local Similarity 62.5%; Pred. No. 1.3e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TNNVLOXT 8  
Db 64 TNNVLOXT 71

RESULT 14  
US-10-157-104-77  
; GENERAL INFORMATION:  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Vernet, Corine A.M  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Tchernev, Velizar T  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Casman, Stacie  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Gerlach, Valerie L  
; APPLICANT: MacDougall, John R.  
; APPLICANT: Stone, David J.  
; APPLICANT: Smithson, Glenda  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 15966-694 CIP2CON2  
; CURRENT APPLICATION NUMBER: US/10/157,104  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: 60/185,674  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR APPLICATION NUMBER: 60/185,535



;; PRIOR FILING DATE: 2000-02-28  
;; PRIOR APPLICATION NUMBER: 60/186,595  
;; PRIOR FILING DATE: 2000-03-03  
;; PRIOR APPLICATION NUMBER: 60/186,604  
;; PRIOR FILING DATE: 2000-03-03  
;; PRIOR APPLICATION NUMBER: 60/186,584  
;; PRIOR FILING DATE: 2000-03-03  
;; PRIOR APPLICATION NUMBER: 60/186,717  
;; PRIOR FILING DATE: 2000-03-03  
;; PRIOR APPLICATION NUMBER: 60/186,716  
;; PRIOR FILING DATE: 2000-03-03  
;; PRIOR APPLICATION NUMBER: 60/186,719  
;; PRIOR FILING DATE: 2000-03-03  
;; PRIOR APPLICATION NUMBER: 60/186,827  
;; PRIOR FILING DATE: 2000-03-03  
;; PRIOR APPLICATION NUMBER: 60/218,323  
;; PRIOR FILING DATE: 2000-07-14  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 104  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 77  
;; LENGTH: 310  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-157-104-77

Query Match 72.2%; Score 26; DB 6; Length 310;  
Best Local Similarity 62.5%; Pred. No. 1.5e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8  
Db 154 SNSVLOST 161

## RESULT 15

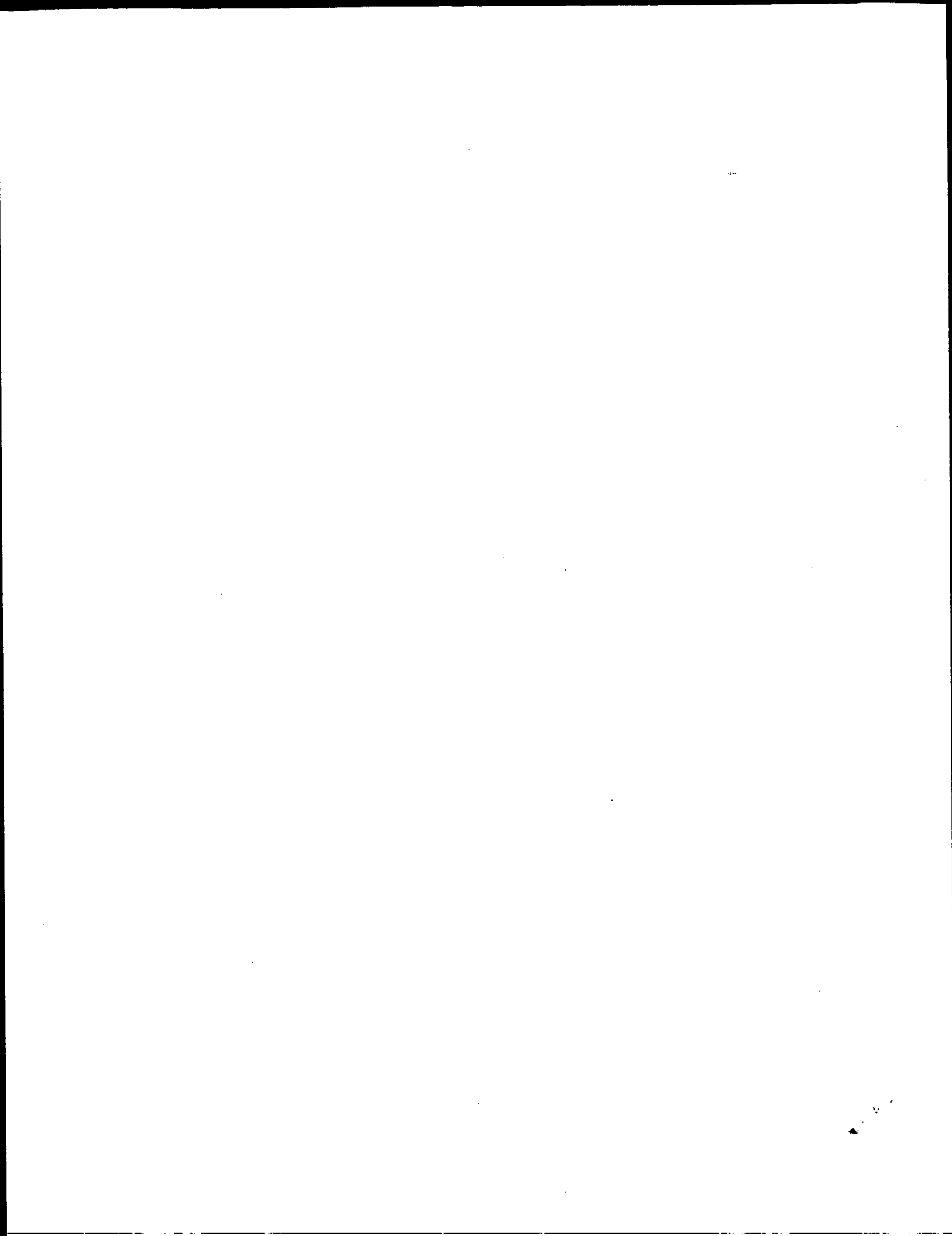
US-10-005-041A-49  
Sequence 49. Application US/10005041A  
GENERAL INFORMATION:  
APPLICANT: Casman, Stacie J  
APPLICANT: Padigar, Muralidhara  
APPLICANT: Burgess, Catherine E  
APPLICANT: Shinkets, Richard A  
APPLICANT: Spytek, Kimberly A  
APPLICANT: Gilbert, Jennifer A  
APPLICANT: Mayotte, Jane E  
APPLICANT: Baumgartner, Jason C  
APPLICANT: Mishra, Vishnu  
APPLICANT: Vernet, Corine AM  
APPLICANT: Dickinson, Kevin S  
APPLICANT: Ballinger, Robert A  
APPLICANT: Wolenc, Adam R  
APPLICANT: Edinger, Shlomit R  
APPLICANT: MacDougall, John R  
APPLICANT: Smithson, Glennnda  
APPLICANT: Ellerman, Karen  
APPLICANT: Stone, David J  
APPLICANT: Gunther, Erik  
APPLICANT: Gerlach, Valerie  
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
FILE REFERENCE: 21402-215  
CURRENT APPLICATION NUMBER: US/10/005,041A  
CURRENT FILING DATE: 2001-12-04  
PRIOR APPLICATION NUMBER: 60/251,459  
PRIOR FILING DATE: 2000-12-05  
PRIOR APPLICATION NUMBER: 60/259,007  
PRIOR FILING DATE: 2000-12-29  
NUMBER OF SEQ ID NOS: 205  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 49  
LENGTH: 310  
TYPE: PRT  
ORGANISM: Homo sapiens

US-10-005-041A-49

Query Match 72.2%; Score 26; DB 6; Length 310;  
Best Local Similarity 62.5%; Pred. No. 1.5e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8  
Db 154 SNSVLOST 161

Search completed: February 26, 2003, 15:42:06  
Job time : 23 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:36:26 ; Search time 30 seconds  
(without alignments)  
7.846 Million cell updates/sec

Title: US-09-476-485A-24  
Perfect score: 36  
Sequence: 1 TNNVLOXT 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA:\*  
1: /cgn2.6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2.6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2.6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2.6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2.6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2.6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result NO.	Score	Query Match %	Length	DB ID	Description
1	34	94.4	8	4	US-08-881-189B-24
2	34	94.4	270	4	US-08-881-189B-2
3	34	94.4	286	4	US-08-881-189B-23
4	29	80.6	1184	4	US-09-541-782-2
5	29	80.6	1184	4	US-09-723-820-2
6	27	75.0	365	1	US-08-089-755A-5
7	27	75.0	365	1	US-08-421-754-5
8	27	75.0	365	2	US-08-421-791-5
9	27	75.0	365	2	US-08-851-088-4
10	27	75.0	365	4	US-08-851-089-10
11	26	72.2	903	4	US-09-228-986-78
12	25	69.4	20	3	US-08-872-094-12
13	25	69.4	188	4	US-08-737-109-14
14	25	69.4	239	2	US-08-665-647-11
15	25	69.4	261	4	US-09-045-973-3
16	25	69.4	299	4	US-09-134-001C-3433
17	25	69.4	307	1	US-08-164-614A-11
18	25	69.4	307	2	US-08-456-489B-11
19	25	69.4	464	2	US-09-021-323-3
20	25	69.4	480	3	US-09-189-035-5
21	25	69.4	480	4	US-09-382-086-5
22	25	69.4	520	4	US-08-961-083-122
23	25	69.4	522	1	US-08-164-614A-10
24	25	69.4	522	2	US-08-456-489B-10
25	25	69.4	536	1	US-08-164-614A-12
26	25	69.4	536	2	US-08-456-489B-12
27	25	69.4	580	1	US-08-309-512-6

28 25 69.4 580 5 PCT-US92-08756A-6 Sequence 6, Appli  
29 25 69.4 599 1 US-08-222-619-2 Sequence 2, Appli  
30 25 69.4 599 4 US-08-221-767-24 Sequence 24, Appli  
31 25 69.4 599 5 PCT-US95-04075-2 Sequence 2, Appli  
32 25 69.4 622 1 US-08-547-197-1 Sequence 1, Appli  
33 25 69.4 622 4 US-08-957-940-1 Sequence 11, Appli  
34 25 69.4 630 4 US-09-228-986-71 Sequence 11, Appli  
35 25 69.4 693 1 US-08-463-620-11 Sequence 71, Appli  
36 25 69.4 693 2 US-08-224-917-11 Sequence 11, Appli  
37 25 69.4 693 2 PCT-US95-03934A-11 Sequence 11, Appli  
38 25 69.4 693 5 US-08-568-459A-8 Sequence 8, Appli  
39 25 69.4 921 2 US-08-487-826B-8 Sequence 8, Appli  
40 25 69.4 921 4 US-09-210-288-8 Sequence 8, Appli  
41 25 69.4 921 3 US-08-872-094-8 Sequence 8, Appli  
42 25 69.4 1040 4 US-08-961-083-118 Sequence 118, Appli  
43 25 69.4 1481 2 US-08-616-844-40 Sequence 40, Appli  
44 25 69.4 1481 2 US-08-599-654-40 Sequence 40, Appli  
45 25 69.4 1481 2

## ALIGNMENTS

RESULT 1  
US-08-881-189B-24  
; Sequence 24, Application US/08881189B  
; Patent No. 6310195  
; GENERAL INFORMATION:  
; APPLICANT: Colucci et al.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED  
; NUMBER OF INVENTION: PROGENITOR CELL PRESERVATION FACTOR  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron, LLP  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11753  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/881,189B  
; FILING DATE: June 24, 1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: 381-44 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-881-189B-24

Query Match 94.4%; Score 34; DB 4; Length 8;  
Best Local Similarity 87.5%; Pred. No. 2e+05;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8

Db 1 TNNVLOXT 8

```
RESULT 2
US-08-881-189B-2
; Sequence 2, Application US/08881189B
; Patent No. 6310195
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; PROGENITOR CELL PRESERVATION FACTOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,189B
; FILING DATE: June 24, 1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-881-189B-2

Query Match 94.4%; Score 34; DB 4; Length 270;
Best Local Similarity 87.5%; Pred. No. 3.8;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
DB 27 TNNVLQXT 34

RESULT 3
US-08-881-189B-23
; Sequence 23, Application US/08881189B
; Patent No. 6310195
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; PROGENITOR CELL PRESERVATION FACTOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
```

```
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,189B
; FILING DATE: June 24, 1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 286 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-881-189B-23
```

```
Query Match 94.4%; Score 34; DB 4; Length 286;
Best Local Similarity 87.5%; Pred. No. 4;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 TNNVLQXT 8
DB 49 TNNVLQXT 56
```

```
RESULT 4
US-09-541-782-2
; Sequence 2, Application US/09541782
; Patent No. 6284480
; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/541,782
; CURRENT FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1184
; TYPE: PRT
; ORGANISM: A. nidulans
US-09-541-782-2
```

```
Query Match 80.6%; Score 29; DB 4; Length 1184;
Best Local Similarity 75.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 TNNVLQXT 8
DB 521 TNDVLQXT 528
```

```
RESULT 5
US-09-723-820-2
; Sequence 2, Application US/09723820
; Patent No. 6468760
; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/723,820
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/541,782
```

;; PRIOR FILING DATE: 2000-04-03  
;; NUMBER OF SEQ ID NOS: 10  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 2  
;; LENGTH: 1184  
;; TYPE: PRT  
;; ORGANISM: A. nidulans  
US-09-723-820-2

Query Match 80.6%; Score 29; DB 4; Length 1184;  
Best Local Similarity 75.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8  
|||  
DB 521 TNDVLQQT 528

RESULT 6  
US-08-089-755A-5  
; Sequence 5, Application US/08089755A  
; Patent No. 5356801  
; GENERAL INFORMATION:  
; APPLICANT: Rambossek, John  
; APPLICANT: Piddington, Chris  
; APPLICANT: Kovacevich, Brian R  
; APPLICANT: Young, Kevin D  
; APPLICANT: Denome, Sylvia A  
; TITLE OF INVENTION: Recombinant DNA Encoding A  
; TITLE OF INVENTION: Desulfurization Biocatalyst  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith and Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02173

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/089,755A  
FILING DATE: 09-JUL-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/911,845  
FILING DATE: 10-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook, David E  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: EBC92-03A  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 365 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-089-755A-5

Query Match 75.0%; Score 27; DB 1; Length 365;  
Best Local Similarity 83.3%; Pred. No. 1.8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
|||  
DB 338 TNNLLQ 343

RESULT 7  
US-08-421-754-5  
; Sequence 5, Application US/08421754  
; Patent No. 5578478  
; GENERAL INFORMATION:  
; APPLICANT: Rambossek, John  
; APPLICANT: Piddington, Chris  
; APPLICANT: Kovacevich, Brian R  
; APPLICANT: Young, Kevin D  
; APPLICANT: Denome, Sylvia A  
; TITLE OF INVENTION: Recombinant DNA Encoding A  
; TITLE OF INVENTION: Desulfurization Biocatalyst  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith and Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/421,754  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/304,081  
FILING DATE: 01-SEP-1994  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/089,755  
FILING DATE: 09-JUL-1993  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/911,845  
FILING DATE: 10-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook, David E  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: EBC92-03AZX  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 365 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-421-754-5

Query Match 75.0%; Score 27; DB 1; Length 365;  
Best Local Similarity 83.3%; Pred. No. 1.8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
|||  
DB 338 TNNLLQ 343

RESULT 8  
US-08-421-791-5  
; Sequence 5, Application US/08421791  
; Patent No. 5879914  
; GENERAL INFORMATION:  
; APPLICANT: Rambossek, John  
; APPLICANT: Piddington, Chris  
; APPLICANT: Kovacevich, Brian R

APPLICANT: Young, Kevin D  
APPLICANT: Denome, Sylvia A  
TITLE OF INVENTION: Recombinant DNA Encoding A  
TITLE OF INVENTION: Desulfurization Biocatalyst  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith and Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/421,791  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/304,081  
FILING DATE: 01-SEP-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/089,755  
FILING DATE: 09-JUL-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/911,845  
FILING DATE: 10-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook, David E  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: EBC92-03A22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 365 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-421-791-5

Query Match 75.0%; Score 27; DB 2; Length 365;  
Best Local Similarity 83.3%; Pred. No. 1.8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
Db 338 TNNLLQ 343  
|||:|

RESULT 9  
US-08-851-088-4  
Sequence 4, Application US/08851088  
Patent No. 5952208  
GENERAL INFORMATION:  
APPLICANT: Darzins, Aldis  
APPLICANT: XI, Lei  
APPLICANT: Childs, John D.  
APPLICANT: Monticello, Daniel J.  
APPLICANT: Squires, Charles H.  
TITLE OF INVENTION: DS2 Gene Expression In Pseudomonas Hosts  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts

COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/851,088  
FILING DATE: 05-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/835,185  
FILING DATE: 07-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmore, Carolyn S.  
REGISTRATION NUMBER: 37,567  
REFERENCE/DOCKET NUMBER: EBC96-06A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781) 861-6240  
TELEFAX: (781) 861-9540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 365 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-851-088-4

Query Match 75.0%; Score 27; DB 2; Length 365;  
Best Local Similarity 83.3%; Pred. No. 1.8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
Db 338 TNNLLQ 343  
|||:|

RESULT 10  
US-08-851-089-10  
Sequence 10, Application US/08851089  
Patent No. 6133016  
GENERAL INFORMATION:  
APPLICANT: Aldis Darzins  
APPLICANT: Gregory T. Mrachko  
TITLE OF INVENTION: A Sphingomonas Biotransformation  
TITLE OF INVENTION: Catalyst  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/851,089  
FILING DATE: 05-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/835,292  
FILING DATE: 07-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmore, Carolyn S.  
REGISTRATION NUMBER: 37,567  
REFERENCE/DOCKET NUMBER: EBC97-06A2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781) 861-6240

TELEFAX: (781) 861-9540  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 365 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-851-089-10

Query Match 75.0%; Score 27; DB 4; Length 365;  
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
 ||:||  
 Db 338 TNNLLQ 343

RESULT 11  
 US-09-228-986-78  
 Sequence 78, Application US/09228986  
 Patent No. 6359198

GENERAL INFORMATION:  
 APPLICANT: Strabala, Timothy  
 APPLICANT: Nieuwenhuizen, Niels  
 TITLE OF INVENTION: Compositions Isolated from Plant Cells  
 TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling  
 FILE REFERENCE: 11000/1020  
 CURRENT APPLICATION NUMBER: US/09/228,986  
 CURRENT FILING DATE: 1999-01-12  
 NUMBER OF SEQ ID NOS: 130  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 78  
 LENGTH: 903  
 TYPE: PRT  
 ORGANISM: Eucalyptus grandis  
 US-09-228-986-78

Query Match 72.2%; Score 26; DB 4; Length 903;  
 Best Local Similarity 71.4%; Pred. No. 7.9e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNVLQXT 8  
 ||:||  
 Db 492 NNLLQGT 498

RESULT 12  
 US-08-872-094-12  
 Sequence 12, Application US/08872094  
 Patent No. 6020128  
 GENERAL INFORMATION:  
 APPLICANT: Steiner, Bret M.  
 TITLE OF INVENTION: DNA Polymerase from Treponema pallidum  
 NUMBER OF SEQUENCES: 14  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Jones & Askew  
 STREET: 191 Peachtree Street, 37th Floor  
 CITY: Atlanta  
 STATE: Georgia  
 COUNTRY: U.S.A.  
 ZIP: 30303  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/872,094  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Warren, William L.

REGISTRATION NUMBER: 36,714  
 REFERENCE/DOCKET NUMBER: 03063-0320  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (404) 818-3700  
 TELEFAX: (404) 818-3799  
 INFORMATION FOR SEQ ID NO: 12:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 20 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 US-08-872-094-12

Query Match 69.4%; Score 25; DB 3; Length 20;  
 Best Local Similarity 62.5%; Pred. No. 19;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8  
 ||:||  
 Db 8 TNDTLQMT 15

RESULT 13  
 US-08-737-109-14  
 Sequence 14, Application US/08737109  
 Patent No. 6455688  
 GENERAL INFORMATION:  
 APPLICANT: SLABAS, Antoni Ryszard  
 APPLICANT: ELBOROUGH, Kieran Michael  
 APPLICANT: BRIGHT, Simon William Jonathan  
 APPLICANT: FENTEN, Philip Anthony  
 TITLE OF INVENTION: Plant Gene Specifying Acetyl Coenzyme A  
 TITLE OF INVENTION: Carboxylase and Transformed Plants Containing Same  
 NUMBER OF SEQUENCES: 32  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.  
 STREET: 1100 New York Avenue, N.W.  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch diskette  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: MS Word  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/737,109  
 FILING DATE: 21-OCT-1996  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/GB94/00846  
 FILING DATE: 02-MAY-1994  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 188 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: YES  
 FRAGMENT TYPE: N-terminal  
 ORIGINAL SOURCE:  
 ORGANISM: Arabidopsis thaliana  
 IMMEDIATE SOURCE:  
 CLONE: pKL081  
 US-08-737-109-14

Query Match 69.4%; Score 25; DB 4; Length 188;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNVL 5  
Db 26 TNNVL 30

## RESULT 14

US-08-665-647-11  
; Sequence 11, Application US/08665647  
; Patent No. 5935803  
; GENERAL INFORMATION:  
; APPLICANT: Dasquez, Nicki J.  
; APPLICANT: Ron, Dorit  
; APPLICANT: Voronova, Anna F.  
; APPLICANT: Napolitano, Eugene W.  
; TITLE OF INVENTION: METHODS TO IDENTIFY IMMUNOMODULATORS  
; TITLE OF INVENTION: USING COGNATE INTERACTION OF PKC-THETA  
; NUMBER OF SEQUENCES: 89  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 2000 Pennsylvania Avenue, NW - Ste. 5500  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1888  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/665,647  
; FILING DATE: 18-JUN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Murashige, Kate H.  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 22550-20025.25  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-1500  
; TELEFAX: (202) 822-0168  
; TELEX: 90-4030 MRSNFOERSWSH  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 239 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-665-647-11

Query Match 69.4%; Score 25; DB 2; Length 239;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NNVLQ 6  
Db 26 NNVLQ 30

## RESULT 15

US-09-045-973-3  
; Sequence 3, Application US/09045973  
; Patent No. 6165767  
; GENERAL INFORMATION:  
; APPLICANT: Lal, Preeti  
; APPLICANT: Yue, Henry  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Baughn, Mariah  
; TITLE OF INVENTION: PROTEIN PHOSPHATASE RELATED MOLECULES  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive

; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/045,973  
; FILING DATE: Filed Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0491 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 261 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: BRAINOT18  
; CLONE: 2534680  
US-09-045-973-3

Query Match 69.4%; Score 25; DB 4; Length 261;  
Best Local Similarity 62.5%; Pred. No. 3.3e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8  
Db 204 TNNILFT 211

Search completed: February 26, 2003, 15:39:11  
Job time : 31 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:48:40 ; Search time 16 Seconds  
(without alignments)  
84.118 Million cell updates/sec

Title: US-09-476-485A-31

Perfect score: 66

Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73:\*

1: piri:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	83.3	105	2 B54864	mannose/glucose-sp
2	43	65.2	249	2 JQ1981	lectin II - Scotch
3	42	63.6	408	2 F63379	conserved hypothet
4	39	59.1	290	2 S63355	lectin-related sto
5	39	59.1	924	2 C83572	organic solvent to
6	38	57.6	467	2 T16319	hypothetical prote
7	38	57.6	705	2 S23352	gene unc-93 protei
8	38	57.6	708	2 T19369	hypothetical prote
9	38	57.6	807	2 T19836	hypothetical prote
10	37	56.1	24	2 S08292	lectin - smooth ra
11	37	56.1	220	2 T39548	hypothetical prote
12	37	56.1	298	2 S77536	lactose transport
13	37	56.1	560	2 S54093	hypothetical prote
14	36	54.5	284	2 T23636	hypothetical prote
15	36	54.5	425	2 C87619	2-isopropylmalate
16	36	54.5	524	2 B70479	hypothetical prote
17	36	54.5	534	2 T30629	hypothetical prote
18	36	54.5	590	2 F75501	hypothetical prote
19	36	54.5	699	2 T33375	sensor histidine k
20	35	53.0	124	2 S67606	hypothetical prote
21	35	53.0	190	2 A55666	probable membrane
22	35	53.0	191	2 JH0605	neurocalcin - frui
23	35	53.0	191	2 A48979	neural visinin-lik
24	35	53.0	193	2 JH0616	visinin-like prote
25	35	53.0	193	2 JH0816	neurocalcin (clone
26	35	53.0	193	2 S47565	neural visinin-lik
27	35	53.0	193	2 S47565	gene Rem-1 protein
28	35	53.0	203	2 T25609	calcium-binding pr
29	35	53.0	266	2 S51833	hypothetical prote
					arcelin-4 precursor

## ALIGNMENTS

### RESULT 1

B54864

mannose/glucose-specific lectin beta chain - Dolichos lab lab (field bean)  
C:Species: Dolichos lab lab (field bean)  
C:Date: 23-Jun-1995 #sequence\_revision 10-Nov-1995 #text\_change 07-May-1999  
C:Accession: B54864; B48591

R:Gowda, L.R.; Savithri, H.S.; Rao, D.R.

J. Biol. Chem. 269, 18789-18793, 1994

A:Title: The complete primary structure of a unique mannose/glucose-specific lectin f

A:Reference number: A54864; MUID:94308133; PMID:8034631

A:Accession: B54864

A:Molecule type: protein

A:Residues: 1-105 <GOW>

C:Keywords: glycoprotein; lectin

Query Match

Best Local Similarity 83.3%; Score 55; DB 2; Length 105;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12

Db 1 AQSLSFXFTKFD 12

### RESULT 2

JQ1981

lectin II - Scotch broom

N:Alternate names: 2-acetamido-2-deoxy-D-galactose binding lectin II

C:Species: Cytisus scoparius (Scotch broom)

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 07-May-1999

C:Accession: JQ1981

R:Konami, Y.; Yamamoto, K.; Osawa, T.; Irimura, T.

J. Biochem. 112, 366-375, 1992

A:Title: The primary structure of the Cytisus scoparius seed lectin and a carbohydrate

A:Reference number: JQ1981; MUID:93054441; PMID:1429525

A:Accession: JQ1981

A:Molecule type: protein

A:Residues: 1-249 <KON>

A:Experimental source: seed

C:Superfamily: plant lectin

C:Keywords: lectin

Query Match

Best Local Similarity 65.2%; Score 43; DB 2; Length 249;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14

Db 1 SEELSFSTKFTD 14

### RESULT 3

F83379 conserved hypothetical protein PA2127 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim  
 ; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path  
 A:Reference number: A82950; MUID:20437337; PMID:10984043  
 A:Accession: F83379  
 A:Molecule type: DNA  
 A:Status: preliminary  
 A:Cross-references: GB:AE004640; GB:AE004091; NID:g9948139; PIDN:AAG05515.1; GSPDB:GN001  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA2127  
 C:Superfamily: Escherichia coli ybdN protein

Query Match 63.6%; Score 42; DB 2; Length 408;  
 Best Local Similarity 57.1%; Pred. No. 3.3;  
 Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 Db 1 AQSLSFXFTKFDLD 14  
 :||| :|||  
 385 ARSLSFGFSQADID 398  
 :||| :|||

RESULT 4  
 S66355 lectin-related storage protein precursor - Cladrasia lutea (fragment)  
 C:Species: Cladrasia lutea  
 C:Date: 19-Mar-1997 #sequence\_revision 29-Aug-1997 #text\_change 20-Aug-1999  
 C:Accession: S66355; S66299  
 R:van Damme, E.J.M.; Barre, A.; Berner, V.; Rouge, P.; van Leuven, F.; Peumans, W.J.  
 Plant Mol. Biol. 29, 579-598, 1995  
 A:Title: A lectin and a lectin-related protein are the two most prominent proteins in th  
 A:Reference number: S66299; MUID:96123235; PMID:8534854  
 A:Accession: S66355  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-290 <VAN>  
 A:Cross-references: EMBL:U21940; NID:g1141754; PIDN:AAC49150.1; PID:g1141755  
 A:Accession: S66299  
 A:Molecule type: protein  
 A:Residues: 37-56 <VAM>  
 A:Experimental source: bark  
 C:Comment: This lectin-related protein has no carbohydrate binding activity.  
 C:Superfamily: plant lectin  
 C:Keywords: lectin  
 F:1-36/Domain: signal sequence (fragment) #status predicted <SIG>  
 F:37-290/Product: lectin-related storage protein #status experimental <MAT>

Query Match 59.1%; Score 39; DB 2; Length 290;  
 Best Local Similarity 70.0%; Pred. No. 8.4;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Db 2 QSLSFXTKFK 11  
 :||| :|||  
 38 EALSFTTKF 47  
 :||| :|||

RESULT 5  
 C83572 organic solvent tolerance protein OstA precursor PA0595 [imported] - Pseudomonas aerugin  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: C83572  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim  
 ; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
 A:Reference number: A82950; MUID:20437337; PMID:10984043  
 A:Accession: C83572  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-924 <STO>  
 A:Cross-references: GB:AE004495; GB:AE004091; NID:g9946459; PIDN:AAG03984.1; GSPDB:GN  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: ostA; PA0595

Query Match 59.1%; Score 39; DB 2; Length 924;  
 Best Local Similarity 50.0%; Pred. No. 29;  
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 SLSFXFTKFDLD 14  
 :||| :|||  
 Db 569 TLKLYTKYDLD 580  
 :||| :|||

RESULT 6  
 T16319 hypothetical protein F41C3.2 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C:Accession: T16319  
 R:Chisoe, S.  
 submitted to the EMBL Data Library, July 1995  
 A:Description: The sequence of C. elegans cosmid F41C3.  
 A:Reference number: Z18494  
 A:Accession: T16319  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-467 <CHI>  
 A:Cross-references: EMBL:U23521; NID:g746536; PID:g746538; PIDN:AAC46809.1; CESP:F41C  
 A:Experimental source: strain Bristol N2  
 C:Genetics:  
 A:Gene: CESP:F41C3.2  
 A:Introns: 40/1; 415/2

Query Match 57.6%; Score 38; DB 2; Length 467;  
 Best Local Similarity 58.3%; Pred. No. 22;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 SLSFXFTKFDLD 14  
 :||| :|||  
 Db 52 SLNFNFSKFEQD 63  
 :||| :|||

RESULT 7  
 S23352 gene unc-93 protein 1 - Caenorhabditis elegans  
 N:Contains: gene unc-93 protein 2  
 C:Species: Caenorhabditis elegans  
 C:Date: 03-Mar-1994 #sequence\_revision 10-Nov-1995 #text\_change 21-Jul-2000  
 C:Accession: S23352; S23353  
 R:Levin, J.Z.; Horvitz, H.R.  
 J. Cell Biol. 117, 143-155, 1992  
 A:Title: The Caenorhabditis elegans unc-93 gene encodes a putative transmembrane prot  
 A:Reference number: S23352; MUID:92210636; PMID:1313436  
 A:Accession: S23352  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-705 <LEV>  
 A:Cross-references: EMBL:X64415; NID:g6909; PIDN:CAA45760.1; PID:g6910  
 A:Accession: S23353  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 6-705 <LE2>  
 A:Cross-references: EMBL:X64415; NID:g6909; PIDN:CAA45761.1; PID:g6911  
 C:Genetics:  
 A:Introns: 36/3; 118/2; 147/2; 215/2; 247/3; 284/2; 329/2; 376/2; 418/2; 449/1; 478/2

F:6-705/Product: gene unc-93 protein 2 #status predicted <MAT2>

Query Match 57.6%; Score 38; DB 2; Length 705;  
Best Local Similarity 70.0%; Pred. No. 33;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 QSLSFXTKFK 11  
||| | |||  
Db 623 QSLQFAFTKY 632

## RESULT 8

T19969 hypothetical protein C46F11.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T19969

R:Burton, J.  
submitted to the EMBL Data Library, November 1996

A:Reference number: Z19206

A:Accession: T19969

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-708 <WIL>

A:Cross-references: EMBL:Z81449; PIDN:CA803760.1; GSPDB:GN000021; CESP:C46F11.1

A:Experimental source: clone C46F11

C:Genetics:

A:Gene: CESP:C46F11.1

A:Map position: 3

A:Introns: 36/3; 118/2; 147/2; 215/2; 247/3; 284/2; 329/2; 369/1; 421/2; 452/1; 481/2; 5

Query Match 57.6%; Score 38; DB 2; Length 708;

Best Local Similarity 70.0%; Pred. No. 33;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 QSLSFXTKFK 11  
||| | |||  
Db 626 QSLQFAFTKY 635

## RESULT 9

T19836 hypothetical protein C38H2.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T19836

R:Sim, M.

submitted to the EMBL Data Library, August 1994

A:Reference number: Z19185

A:Accession: T19836

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-807 <WIL>

A:Cross-references: EMBL:Z35641; PIDN:CAA84706.1; GSPDB:GN000021; CESP:C38H2.1

A:Experimental source: clone C38H2

C:Genetics:

A:Gene: CESP:C38H2.1

A:Map position: 3

A:Introns: 42/2; 81/3; 190/2; 215/1; 257/3; 270/1; 388/1; 443/2; 461/2; 528/3; 571/3; 61

Query Match 57.6%; Score 38; DB 2; Length 807;

Best Local Similarity 64.3%; Pred. No. 38;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AQLSFXTKFDLD 14  
||| | | |  
Db 704 AQLSFAFRWFLD 717

## RESULT 10

S08292

lectin - smooth rattlebox (fragment)

C:Species: Crotalaria pallida (smooth rattlebox)

C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 18-Jun-1993  
C:Accession: S08292

R:Nguyen, O.K.; Guillaume, J.L.; Hoebeke, J.

Biochim. Biophys. Acta 1033, 210-213, 1990

A:Title: A blood group A specific lectin from the seeds of *Crotalaria striata*.

A:Reference number: S08292; MUID:90167102; PMID:2306467

A:Accession: S08292

A:Molecule type: protein

A:Residues: 1-24 <NGO>

A:Note: the source is designated as *Crotalaria striata*

C:Keywords: lectin

Query Match 56.1%; Score 37; DB 2; Length 24;

Best Local Similarity 70.0%; Pred. No. 1.5;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXTKFDLD 14  
|| | ||| |  
Db 5 SFSTKPFSTD 14

## RESULT 11

T39548

hypothetical protein SPBC16A3.04 - fission yeast (*Schizosaccharomyces pombe*)

C:Species: *Schizosaccharomyces pombe*

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: T39548

R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.

submitted to the EMBL Data Library, February 1998

A:Reference number: Z21862

A:Accession: T39548

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-220 <WOO>

A:Cross-references: EMBL:AL021748; PIDN:CAAL6855.1; GSPDB:GN000067; SPDB:SPBC16A3.04

A:Experimental source: strain 972h-; cosmid c16A3

C:Genetics:

A:Gene: SPDB:SPBC16A3.04

A:Map position: 2

Query Match 56.1%; Score 37; DB 2; Length 220;

Best Local Similarity 57.1%; Pred. No. 15;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 AQLSFXTKFDLD 14  
|| | ||| |  
Db 167 AQLGAVTKSDLE 180

## RESULT 12

S77536

lactose transport system permease protein lacF - *Synechocystis* sp. (strain PCC 6803)

N:Alternate names: protein slr1202

C:Species: *Synechocystis* sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000

C:Accession: S77536

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,

O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocys*

s.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S77536

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-298 <KAN>

A:Cross-references: EMBL:D90905; GB:AB001339; NID:g1652360; PIDN:BAAL7383.1; PID:g165

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Gene: lacF

C:Superfamily: inner membrane protein ugpA

C:Keywords: binding protein-dependent transport system; lactose transport; membrane p

Query Match 56.1%; Score 37; DB 2; Length 298;  
Best Local Similarity 46.2%; Pred. No. 21;  
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 QSLSFXTKFDLD 14  
II: I II::II  
DB 37 QAFSLSFYQYELD 49

RESULT 13  
S54093  
hypothetical protein YPR072w - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein YP9499.27  
C:Species: Saccharomyces cerevisiae  
C>Date: 08-Jul-1995 #sequence\_revision 19-Oct-1995 #text\_change 19-Apr-2002  
C:Accession: S54093; S69060  
R:Badcock, K.; Churcher, C.M.  
submitted to the EMBL Data Library, May 1995  
A:Reference number: S54059  
A:Accession: S54093  
A:Molecule type: DNA  
A:Residues: 1-560 <BAD>  
A:Cross-references: EMBL:Z49219; NID:g805025; PIDN:CAA89189.1; PID:g805052; MIPS:YPR072w  
A:Experimental source: Strain AB972  
R:Couch, J.  
submitted to the EMBL Data Library, March 1996  
A:Description: The sequence of S. cerevisiae cosmid 9513.  
A:Reference number: S69057  
A:Accession: S69060  
A:Molecule type: DNA  
A:Residues: 1-560 <COU>  
A:Cross-references: EMBL:U51033; NID:g1230676; PIDN:AAB68123.1; PID:g1230680; MIPS:YPR072w  
C:Genetics:  
A:Gene: SGD:NOT5  
A:Cross-references: SGD:S0006276  
A:Map position: 16R  
C:Superfamily: hypothetical protein YPR072w

Query Match 56.1%; Score 37; DB 2; Length 560;  
Best Local Similarity 66.7%; Pred. No. 40;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDLD 14  
II: I II::II  
DB 464 SLARIPMKFDLD 475

RESULT 14  
T23636  
hypothetical protein M01B2.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 26-May-2000  
C:Accession: T23636  
R:Lloyd, C.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: Z19776  
A:Accession: T23636  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-284 <WIL>  
A:Cross-references: EMBL:Z83116; PIDN:CAB05561.1; GSPDB:GN00023; CBSP:M01B2.3  
A:Experimental source: clone M01B2  
C:Genetics:  
A:Gene: CBSP:M01B2.3  
A:Map position: 5  
A:Introns: 211/3  
C:Superfamily: Caenorhabditis elegans hypothetical protein B0250.6

Query Match 54.5%; Score 36; DB 2; Length 284;  
Best Local Similarity 46.2%; Pred. No. 31;  
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 QSLSFXTKFDLD 14  
II: I II::II  
DB 40 QSMTFIYWKFSVD 52

RESULT 15  
C87619  
hypothetical protein CC2989 [Imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: C87619  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko, N. J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: C87619  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-425 <STO>  
A:Cross-references: GB:AE005673; NID:gl3424623; PIDN:AAK24951.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC2989

Query Match 54.5%; Score 36; DB 2; Length 425;  
Best Local Similarity 63.6%; Pred. No. 47;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AOSLSFXFTKF 11  
II: I II::II  
DB 361 AOALSFVFTPY 371

Search completed: February 26, 2003, 14:49:22  
Job time : 17 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:48:40 ; Search time 11 Seconds  
(without alignments)  
52.788 Million cell updates/sec

Title: US-09-476-485A-31  
Perfect score: 66  
Sequence: 1 AQSLSFXTKFDL14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	83.3	237	1	LECA_DOLLA
2	43	65.2	248	1	LEC2_CYTSC
3	39	59.1	290	1	LECR_CLALU
4	37	56.1	24	1	LEC_CROST
5	37	56.1	560	1	NOT5_YEAST
6	36	54.5	520	1	LEU1_AQUAE
7	35	53.0	189	1	NCAB_DRONE
8	35	53.0	190	1	VIS1_HUMAN
9	35	53.0	192	1	NCAD_CHICK
10	35	53.0	192	1	NCAD_HUMAN
11	35	53.0	192	1	NCAD_MOUSE
12	35	53.0	192	1	VIS3_CHICK
13	35	53.0	192	1	VIS3_HUMAN
14	35	53.0	192	1	VIS3_MOUSE
15	35	53.0	236	1	LECA_CRAFL
16	35	53.0	240	1	LECS_VATWA
17	35	53.0	266	1	ARC4_PHAVU
18	35	53.0	431	1	EF1G_DRONE
19	35	53.0	847	1	YAF6_YEAST
20	35	53.0	930	1	PMP8_CHLPN
21	34	51.5	165	1	NCAB_BOVIN
22	34	51.5	174	1	NCAB_NEUCR
23	34	51.5	185	1	LEC_VICVI
24	34	51.5	190	1	VIS2_RAT
25	34	51.5	192	1	HIPP_HUMAN
26	34	51.5	192	1	NECX_APLCA
27	34	51.5	196	1	CAGS_HELPY
28	34	51.5	199	1	CAGS_HELPJ
29	34	51.5	243	1	LEC2_CYTSE
30	34	51.5	270	1	LECB_SOPJA
31	34	51.5	285	1	LCB1_ROBPS
32	34	51.5	285	1	LCST1_ROBPS
33	34	51.5	285	1	LCS2_ROBPS

34	34	51.5	286	1	LCB2_ROBPS	Q42372 robinia pse
35	34	51.5	292	1	LECS_SOPJA	P93535 sophora jap
36	34	51.5	363	1	YRU5_CAEEL	Q09351 caenorhabdi
37	34	51.5	435	1	YCW9_YEAST	P25646 saccharomyc
38	34	51.5	449	1	PCOI_HUMAN	Q15113 homo sapien
39	34	51.5	575	1	11OR_MOUSE	O61216 mus musculu
40	34	51.5	706	1	MR11_MOUSE	O61216 mus musculu
41	34	51.5	706	1	MR11_MOUSE	O61216 mus musculu
42	34	51.5	708	1	MR11_MOUSE	O61216 mus musculu
43	34	51.5	708	1	MR11_MOUSE	O61216 mus musculu
44	33	50.0	15	1	LEC2_PSOSC	P49959 rattus norv
45	33	50.0	24	1	LEC2_CROJUF	P22585 psophocarpu
	33	50.0	167	1	YMH6_CAEEL	P16352 crotalaria
	33	50.0				P34473 caenorhabdi

## ALIGNMENTS

RESULT 1  
LECA\_DOLLA STANDARD; PRT; 237 AA.  
AC P38662;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE Lectin.  
OS Dolichos lab lab (Field bean).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Dolichos.  
OX NCBI\_TaxID=35936;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=cv. Lignosus; TISSUE=Seed;  
RX MEDLINE=94308133; PubMed=8034631;  
RA Gowda L.R., Savithri H.S., Rajagopal Rao D.;  
RT "The complete primary structure of a unique mannose/glucose-specific  
lectin from field bean (Dolichos lab lab)".  
RL J. Biol. Chem. 269:18789-18793(1994).  
CC -1- FUNCTION: D-MANNOSE/D-GLUCOSE-BINDING LECTIN. REQUIRES CA2+ AND  
CC MN2+ IONS FOR FULL ACTIVITY.  
CC -1- SUBUNIT: Tetramer of two alpha and two beta chains.  
CC -1- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.  
DR HSP; P02866; IONA.  
DR InterPro; IPR000985; Lectin\_legA.  
DR InterPro; IPR001220; Lectin\_legB.  
DR Pfam; PF00138; lectin\_legA; 1.  
DR Pfam; PF00139; lectin\_legB; 1.  
DR ProDom; PD000671; Lectin\_legA; 1.  
DR ProDom; PD000711; Lectin\_legB; 1.  
DR PROSITE; PS00307; LECTIN\_LEGUME\_ALPHA; 1.  
DR PROSITE; PS00308; LECTIN\_LEGUME\_BETA; 1.  
KW Lectin; Calcium; Manganese; Glycoprotein; Acetylation.  
FT CHAIN 1 105  
FT CARBOHYD 106 237  
FT MOD\_RES 69 69  
FT N-LINKED (GLCNAC...) (POTENTIAL).  
FT METAL 115 115  
FT METAL 117 117  
FT METAL 120 120  
FT METAL 122 122  
FT METAL 127 127  
FT METAL 132 132  
SQ SEQUENCE 237 AA; 25718 MW; BEB7E84DC2895327 CRC64;

Query Match 83.3%; Score 55; DB 1; Length 237;  
Best Local Similarity 91.7%; Pred. No. 0.0025;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXTKFD 12  
DB 1 AQSLSFXTKFD 12

```

RESULT 2
LEC2_CYTSC
ID LEC2_CYTSC STANDARD; PRT; 248 AA.
AC P29257;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE 2-acetamido-2-deoxy-D-galactose-binding seed lectin II (CSII).
OS Cytisus scoparius (Scotch broom).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Cytisus.
OX NCBI_TaxID=3835;
RN [1]
RP SEQUENCE.
RC TISSUE=Seed;
RX MEDLINE=93054441; PubMed=1429525;
RA Konami Y., Yamamoto K., Osawa T., Irimura T.;
RT "The primary structure of the Cytisus scoparius seed lectin and a
RT carbohydrate-binding peptide.";
RL J. Biochem. 112:366-375(1992).
RN [2]
RP SEQUENCE OF 1-32.
RX MEDLINE=84307428; PubMed=6477513;
RA Young N.M., Watson D.C., Williams R.E.;
RT "Structural differences between two lectins from Cytisus scoparius,
RT both specific for D-galactose and N-acetyl-D-galactosamine.";
RL Biochem. J. 222:41-48(1984).
CC -1- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
KW HSSP; P19588; 1LUL.
DR InterPro: IPR000985; Lectin_legA.
DR InterPro: IPR001220; Lectin_legB.
DR Pfam: PF00138; lectin_legA; 1.
DR Pfam: PF00139; lectin_legB; 1.
DR ProDom: PD000671; Lectin_legA; 1.
DR ProDom: PD000711; Lectin_legB; 1.
DR PROSITE: PS00307; LECTIN_LEGUME_BETA; 1.
DR PROSITE: PS00308; LECTIN_LEGUME_ALPHA; 1.
CC -1- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
KW HSSP; P19588; 1LUL.
DR InterPro: IPR000985; Lectin_legA.
DR InterPro: IPR001220; Lectin_legB.
DR Pfam: PF00138; lectin_legA; 1.
DR Pfam: PF00139; lectin_legB; 1.
DR ProDom: PD000671; Lectin_legA; 1.
DR ProDom: PD000711; Lectin_legB; 1.
DR PROSITE: PS00307; LECTIN_LEGUME_BETA; 1.
DR PROSITE: PS00308; LECTIN_LEGUME_ALPHA; 1.
KW Lectin; Calcium; Manganese; Glycoprotein.
FT METAL 128 128 MANGANESE (BY SIMILARITY).
FT METAL 130 130 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 132 132 CALCIUM (BY SIMILARITY).
FT METAL 134 134 CALCIUM (BY SIMILARITY).
FT METAL 138 138 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 144 144 MANGANESE (BY SIMILARITY).
FT METAL 148 148 N-LINKED (GLCNAC... ) (PARTIAL).
FT CARBOHYD 119 119
SQ SEQUENCE 248 AA; 27028 MW; C960F457C8C9F62A CRC64;

Query Match 65.2%; Score 43; DB 1; Length 248;
Best Local Similarity 57.1%; Pred. No. 0.5;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14
DB 1 SEELSFSTKFKTD 14

RESULT 3
LECR_CLALU
ID LECR_CLALU STANDARD; PRT; 290 AA.
AC Q39527;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Lectin-related protein precursor (CLLRP) (LRPCL) (Fragment).
OS Cladrasia lutea (Yellow wood).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Cladrasia.
OX NCBI_TaxID=38412;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 37-56.
RC TISSUE=Bark;

```

```

RX MEDLINE=96123235; PubMed=8534854;
RA van Damme E.J.M., Barre A., Bemer V., Rouge P., van Leuven F.,
RA Peumans W.J.;
RT "A lectin and a lectin-related protein are the two most prominent
RT proteins in the bark of yellow wood (Cladrasia lutea).";
RL Plant Mol. Biol. 29:579-598(1995).
CC -1- FUNCTION: DOES NOT HAVE ANY CARBOHYDRATE BINDING OR AGGLUTINATION
CC ACTIVITY.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: TO OTHER LEGUMINOUS LECTINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; U21940; AAC49150.1; -.
DR HSSP; P19588; 1LUL.
DR InterPro: IPR000985; Lectin_legA.
DR InterPro: IPR001220; Lectin_legB.
DR Pfam: PF00138; lectin_legA; 1.
DR Pfam: PF00139; lectin_legB; 1.
DR ProDom: PD000671; Lectin_legA; 1.
DR ProDom: PD000711; Lectin_legB; 1.
DR PROSITE: PS00307; LECTIN_LEGUME_BETA; 1.
DR PROSITE: PS00308; LECTIN_LEGUME_ALPHA; 1.
KW Lectin; Calcium; Manganese; Glycoprotein; Signal.
FT NON_TER 1 1
FT SIGNAL <1 36 LECTIN-RELATED PROTEIN.
FT CHAIN 37 290 MANGANESE (BY SIMILARITY).
FT METAL 161 161 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 163 163 CALCIUM (BY SIMILARITY).
FT METAL 167 167 CALCIUM (BY SIMILARITY).
FT METAL 170 170 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 175 175 MANGANESE (BY SIMILARITY).
FT CARBOHYD 119 119 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 147 147 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 152 152 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 290 AA; 31378 MW; 64F2DBE7B2E20B14 CRC64;

Query Match 59.1%; Score 39; DB 1; Length 290;
Best Local Similarity 70.0%; Pred. No. 3.3;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 QLSLFXFTKF 11
DB 38 EALSFTTFKF 47

RESULT 4
LEC_CROST
ID LEC_CROST STANDARD; PRT; 24 AA.
AC P16351;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-MAY-1992 (Rel. 22, Last annotation update)
DE Lectin (Fragment).
OS Crotalaria striata (Smooth rattlebox) (Crotalaria pallida).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Crotalariaeae;
OC Crotalaria.
OX NCBI_TaxID=3830;
RN [1]
RP SEQUENCE.
RX MEDLINE=90167102; PubMed=2306467;
RA Khang N.O., Guillaume J.-L., Hoebeke J.;
RT "A blood group A specific lectin from the seeds of Crotalaria
RT striata.";

```

RL Biochim. Biophys. Acta 1033:210-213(1990).  
 CC -!- FUNCTION: AGGLUTINATES ERYTHROCYTES OF BLOOD GROUP A. BINDS IN  
 CC DECREASING ORDER OF AFFINITY: N-ACETYL-D-GALACTOSAMINE,  
 CC D-GALACTOSE, AND D-GALACTOSAMINE.  
 CC  
 CC -!- SUBUNIT: HOMOTETRAMER.  
 CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.  
 CC PIR: S08292; S08292.  
 DR InterPro: IPR000985; Lectin\_legA.  
 DR InterPro: IPR001220; Lectin\_legB.  
 DR Pfam: PF00139; lectin\_legB; 1.  
 DR PROSITE: PS00307; LECTIN\_LEGUME\_BETA; PARTIAL.  
 DR PROSITE: PS00308; LECTIN\_LEGUME\_ALPHA; PARTIAL.  
 KW Lectin; Glycoprotein.  
 FT NON\_TER 24  
 SQ SEQUENCE 24 AA; 2875 MW; DC62B82CD9F9BB66 CRC64;  
  
 Query Match 56.1%; Score 37; DB 1; Length 24;  
 Best Local Similarity 70.0%; Pred. No. 0.58;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
 Qy 5 SPXFTKFDLD 14  
 |||||  
 Db 5 SFSFTKFDST 14  
  
 RESULT 5  
 NOT5\_YEAST  
 ID NOT5\_YEAST STANDARD; PRT; 560 AA.  
 AC Q12514;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE General negative regulator of transcription subunit 5.  
 GN NOT5 OR YPR072W OR YP9499.27.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 RN NCBI\_TaxID=4932;  
 FX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c / AB972;  
 RA Cough J.;  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c / AB972;  
 RX MEDLINE=97313271; PubMed=9169875;  
 RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansong W.,  
 RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,  
 RA Borstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,  
 RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,  
 RA Dietrich F.S., Dilius H., DiPaolo T., Dubois E., Duesterhoeft A.,  
 RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,  
 RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,  
 RA Hunnicke-Smith S., Hyman R., Johnston M., Kaiman S., Kleine K.,  
 RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,  
 RA Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,  
 RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,  
 RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,  
 RA Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,  
 RA Urrestarazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,  
 RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,  
 RA Zhong W.W., Zollner A., Vo D.H., Hani J.;  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";  
 RN Nature 387:103-105(1997).  
 FX [3]  
 RP CHARACTERIZATION.  
 RX MEDLINE=98172735; PubMed=9511744;  
 RA Oberholzer U., Collart M.A.;  
 RT "Characterization of NOT5 that encodes a new component of the Not  
 RT protein complex."  
 RL Gene 207:61-69(1998).  
 CC -!- FUNCTION: NEGATIVELY REGULATES THE BASAL AND ACTIVATED

CC TRANSCRIPTION OF MANY GENES. PREFERENTIALLY AFFECTS TC-TYPE TATA  
 CC ELEMENT-DEPENDENT TRANSCRIPTION. COULD DIRECTLY OR INDIRECTLY  
 CC INHIBIT COMPONENT(S) OF THE GENERAL TRANSCRIPTION MACHINERY.  
 CC -!- SUBUNIT: FORMS A COMPLEX THAT COMPRISE NOT1, NOT2, NOT3, NOT4 AND  
 CC NOT5.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- SIMILARITY: TO YEAST NOT3.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U51033; AAB68123.1; -;  
 CC EMBL; Z71255; CAA94980.1; -;  
 CC EMBL; Z49219; CAA89189.1; -;  
 CC TRANSFAC: T03515; -;  
 CC SGD; S0006276; YPR072W.  
 KW Nuclear protein; Transcription regulation; Repressor; Coiled coil.  
 FT DOMAIN 3 26 COILED COIL (POTENTIAL).  
 FT DOMAIN 37 71 COILED COIL (POTENTIAL).  
 FT DOMAIN 124 177 COILED COIL (POTENTIAL).  
 SQ SEQUENCE 560 AA; 65854 MW; 07313711983E92BC CRC64;  
  
 Query Match 56.1%; Score 37; DB 1; Length 560;  
 Best Local Similarity 66.7%; Pred. No. 16;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
 Qy 3 SLSPXFTKFDLD 14  
 |||||  
 Db 464 SLRIFMKFDLD 475  
  
 RESULT 6  
 LEU1\_AQUAE  
 ID LEU1\_AQUAE STANDARD; PRT; 520 AA.  
 AC O67862;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 2-isopropylmalate synthase (EC 4.1.3.12) (alpha-isopropylmalate  
 DE synthase) (Alpha-IPM synthetase).  
 GN LEUA OR AQ\_2090  
 OS Aquifex aeolicus.  
 OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;  
 CC Aquifex.  
 CX NCBI\_TaxID=63363;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VF5;  
 RX MEDLINE=98196666; PubMed=9537320;  
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
 RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aufay M., Huber R.,  
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 RT aeolicus".  
 RL Nature 392:353-358(1998).  
 CC -!- FUNCTION: Catalyzes the condensation of the acetyl group of  
 CC acetyl-CoA with 3-methyl-2-oxobutanoate (2-oxoisovalerate) to form  
 CC 3-carboxy-3-hydroxy-4-methylpentanoate (2-isopropylmalate).  
 CC -!- CATALYTIC ACTIVITY: 3-carboxy-3-hydroxy-4-methylpentanoate + H(2)O.  
 CC acetyl-CoA + 3-methyl-2-oxobutanoate + H(2)O.  
 CC -!- PATHWAY: Leucine biosynthesis; first step.  
 CC -!- SUBUNIT: Homotetramer (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE ALPHA-IPM SYNTHETASE / HOMOCITRATE  
 CC SYNTHASE FAMILY. LEUA 1 SUBFAMILY.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AE000772; AAC07824.1; ALT\_INIT.  
 DR InterPro; IPR002034; AIPM/Hcgt\_synth.  
 DR InterPro; IPR000891; HMG1-like.  
 DR Pfam; PF00682; HMG1-like; 1.  
 DR TIGRFAMs; TIGR00973; leuA\_bact; 1.  
 DR PROSITE; PS00815; AIPM\_HOMOCIT\_SYNTH\_1; 1.  
 DR PROSITE; PS00816; AIPM\_HOMOCIT\_SYNTH\_2; 1.  
 KW Leucine biosynthesis; Lyase; Complete proteome.  
 SQ SEQUENCE 520 AA; 58103 MW; 95B210FA872871D4 CRC64;

Query Match 54.5%; Score 36; DB 1; Length 520;  
 Best Local Similarity 53.8%; Pred. No. 23;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 QSLSPXFTKFDLD 14  
 : | | | | : ||  
 Db 343 EELGFKFTKEELD 355

## RESULT 7

ID NCAH\_DROME STANDARD; PRT; 189 AA.  
 AC P42325; Q9Y967;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Neurocalcin homolog.  
 GN NCA OR CG7641.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Canton-S;  
 RX MEDLINE=95081147; PubMed=7989365;  
 RA Teng D.H.-F., Chen C.-K., Hurley J.B.;  
 RT "A highly conserved homologue of bovine neurocalcin in Drosophila  
 RT melanogaster is a Ca(2+)-binding protein expressed in neuronal  
 RT tissues.";  
 RL J. Biol. Chem. 269:31900-31907(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Fendell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [3]  
 RP MYRISTOYLATION, AND FUNCTION.  
 RX MEDLINE=96215323; PubMed=8626592;  
 RA Faurobert E., Chen C.-K., Hurley J.B., Teng D.H.-F.;  
 RT "Drosophila neurocalcin, a fatty acylated, Ca2+-binding protein that  
 RT associates with membranes and inhibits in vitro phosphorylation of  
 RT bovine rhodopsin.";  
 RL J. Biol. Chem. 271:10256-10262(1996).  
 CC -!- FUNCTION: INHIBITS THE PHOSPHORYLATION OF RHODOPSIN IN A CALCIUM-  
 CC DEPENDENT MANNER. PROBABLY BINDS TWO OR THREE CALCIUM IONS.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN NEURONAL TISSUES. HIGH LEVEL  
 CC EXPRESSION SEEN IN THE CORTICAL REGIONS OF THE CENTRAL BRAIN AND  
 CC LOWER LEVELS IN THE LAMINA, THE FIRST OPTIC LOBE OF THE BRAIN. IT  
 CC IS ALSO FOUND IN THE THORACIC GANGLIA.  
 CC -!- DEVELOPMENTAL STAGE: FOUND IN THE EMBRYOS, LARVAE AND PUPAE.  
 CC EXPRESSION IN THE ADULT HEADS IS HIGHER THAN IN THE BODIES.  
 CC -!- MASS SPECTROMETRY: MW=21975.51; METHOD=Electrospray.  
 CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BELONGS TO  
 CC THE RECOVERIN SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; U15735; AAA62152.1; -.  
 DR EMBL; AE003515; AAF49082.1; -.  
 DR HSSP; P36610; 1G81.

DR FlyBase; FBgn0013303; Nca.  
 DR InterPro; IPR002048; EF-hand.  
 DR InterPro; IPR001125; Recoverin.  
 DR Pfam; PF00036; efhand; 3.

DR PRINTS; PR00450; RECOVERIN.

DR ProDom; PD000012; EF-hand; 1.

DR SMART; SM000054; EFh; 3.

DR PROSITE; PS00018; EF\_HAND; 3.

KW Calcium-binding; Repeat; Myristate.

FT INIT\_MET 0 0

FT LIPID 1 1 MYRISTATE.

FT DOMAIN 35 46 ANCESTRAL CALCIUM SITE 1.

FT CA\_BIND 72 83 EF-HAND 2 (POTENTIAL).

FT CA\_BIND 108 119 EF-HAND 3 (POTENTIAL).

FT CA\_BIND 156 167 EF-HAND 4 (POTENTIAL).

SQ SEQUENCE 189 AA; 21762 MW; 3715201BEA2824F CRC64;

Query Match 53.0%; Score 35; DB 1; Length 189;

Best Local Similarity 46.2%; Pred. No. 12;

Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 QSLSPXFTKFDLD 14

: | | | | : |||

Db 98 QKLKWAFFSMYDLD 110



```

RESULT 8
VISL_HUMAN STANDARD; PRT; 190 AA.
AC 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 28, Last sequence update)
DT 15-FEB-1994 (Rel. 41, Last annotation update)
DE VISLIP-like protein 1 (VILIP-1) (Neural visinin-like protein 1) (NVL-
DE 1) (NVP-1) (21 kDa CABP) (Neurocalcin alpha) (Hippocalcin-like protein
DE 3) (HLP3).
GN VSNL1 OR VISL1.
OS Homo sapiens (Human),
OS Mus musculus (Mouse),
OS Rattus norvegicus (Rat),
OS Bos taurus (Bovine), and
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606, 10090, 10116, 9913, 9031;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Human;
RX MEDLINE=96079121; PubMed=8530085;
RA Polymopoulos M.H., Ide S., Soares M.B., Lennon G.G.;
RT "Sequence characterization and genetic mapping of the human VSNL1
RT gene, a homologue of the rat visinin-like peptide RNVPL1.";
RL Genomics 29:273-275(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Human;
RA Bellingham J.;
RT "Peptide conservation between avian and mammalian visinin-like
RT proteins.";
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=Human;
RX MEDLINE=99449057; PubMed=10520747;
RA Kobayashi M., Sakai E., Furuta Y., Takamatsu K.;
RT "Isolation of two human cDNAs, HLP3 and HLP4, homologous to the
RT neuron-specific calcium-binding protein genes.";
RL DNA Seq. 9:171-176(1998).
RN [4]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC SPECIES=Rat; TISSUE=Brain;
RX MEDLINE=92272711; PubMed=1375457;
RA Kuno T., Kajimoto Y., Hashimoto T., Mukai H., Shirai Y., Saheki S.,
RA Tanaka C.;
RT "cDNA cloning of a neural visinin-like Ca(2+)-binding protein.";
RL Biochem. Biophys. Res. Commun. 184:1219-1225(1992).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=Mouse; STRAIN=BALB/C;
RA Kuno T.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES=Bovine; TISSUE=Brain;
RX MEDLINE=98228226; PubMed=9560316;
RA Kato M., Watanabe Y., Iino S., Takaoka Y., Kobayashi S., Haga T.,
RA Hidaka H.;
RT "Cloning and expression of a cDNA encoding a new neurocalcin isoform
RT (neurocalcin alpha) from bovine brain.";
RL Biochem. J. 331:871-876(1998).
RN [7]
RP SEQUENCE FROM N.A.
RC SPECIES=Chicken; STRAIN=White leghorn; TISSUE=Brain;
RX MEDLINE=93061804; PubMed=1359372;
RA Lenz S., Henschel Y., Zopf D., Voss B., Gundelfinger E.D.;
RT "VILIP, a cognate protein of the retinal calcium binding proteins
RT visinin and recoverin, is expressed in the developing chicken
RT brain.";
RL Brain Res. Mol. Brain Res. 15:133-140(1992).

```

---

```

RN [8]
RP SEQUENCE OF 7-26 AND 99-129.
RC SPECIES=Bovine; TISSUE=Brain;
RX MEDLINE=92406915; PubMed=1527077;
RA Terasawa M., Nakano A., Kobayashi R., Hidaka H.;
RT "Neurocalcin: a novel calcium-binding protein from bovine brain.";
RL J. Biol. Chem. 267:19596-19599(1992).
CC -1- FUNCTION: REGULATES (IN VITRO) THE INHIBITION OF RHODOPSIN
CC PHOSPHORYLATION IN A CALCIUM-DEPENDENT MANNER.
CC -1- TISSUE SPECIFICITY: BRAIN AND RETINA. NEURON-SPECIFIC IN THE
CC CENTRAL AND PERIPHERAL NERVOUS SYSTEM.
CC -1- MISCELLANEOUS: PROBABLY BINDS THREE CALCIUM IONS.
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BELONGS TO
CC THE RECOVERIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U14747; AAA91295.1; -
CC EMBL; AF039555; AAD02174.1; -
CC EMBL; AB001104; BAA86891.1; -
CC EMBL; D10666; BAA01517.1; -
CC EMBL; D21165; BAA04701.1; -
CC EMBL; AB006006; BAA28716.1; -
CC EMBL; X63530; CAA45093.1; -
CC PIR; JH0605; JH0605.
CC PIR; S18904; S18904.
CC PIR; A48979; A48979.
CC HSSP; P36610; Ic81.
CC Genew; HGNC:12722; VSNL1.
CC MIM; 600817; -
CC MGI; 1349453; Vsnll.
CC InterPro; IPR002048; EF-hand.
CC InterPro; IPR001125; Recoverin.
CC Pfam; PF00036; efhand; 6.
CC PRINTS; PR00450; RECOVERIN.
CC ProDom; PD000012; EF-hand; 1.
CC SMART; SM00054; EFh; 3.
CC PROSITE; PS00018; EF_HAND; 3.
CC Calcium-binding; Repeat; Myristate.
FT INIT_MET 0 0 BY SIMILARITY.
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
FT DOMAIN 35 46 ANCESTRAL CALCIUM SITE 1 (POTENTIAL).
FT CA_BIND 72 83 EF-HAND 2 (POTENTIAL).
FT CA_BIND 108 119 EF-HAND 3 (POTENTIAL).
FT CA_BIND 158 169 EF-HAND 4 (POTENTIAL).
FT CONFLICT 24 24 H -> A (IN REF. 8).
FT CONFLICT 64 64 A -> G (IN REF. 1).
FT CONFLICT 74 74 N -> I (IN REF. 1).
FT CONFLICT 117 117 R -> P (IN REF. 8).
FT CONFLICT 139 139 M -> K (IN REF. 5).
FT CONFLICT 157 157 M -> K (IN REF. 3).
FT CONFLICT 168 168 D -> G (IN REF. 1).
FT CONFLICT 171 171 K -> R (IN REF. 1).
SQ SEQUENCE 190 AA; 22011 MW; 954A3BE018A149D8 CRC64;
Query Match 53.0%; Score 35; DB 1; Length 190;
Best Local Similarity 46.2%; Pred. No. 12;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 2 QSLSFXTKFDLD 14
| | | |
Db 98 QKLNWAFNMYDLD 110
RESULT 9
NCAD_CHICK STANDARD; PRT; 192 AA.
ID NCAD_CHICK

```

```

AC 012953;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Neurocalcin delta.
GN NCALD.
OS Gallus gallus (Chicken), and
OS Poephila guttata (zebra finch) (Taeniopygia guttata).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031, 59729;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Chicken; TISSUE=Gizzard;
RA Schonekess B.O., Walsh M.P.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=P.guttata; TISSUE=Telencephalon;
RA Wade J., Peabody C.F., Smith G.W.;
RT "Sexually dimorphic pattern of neurocalcin expression in the
RT developing zebra finch telencephalon.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY BE INVOLVED IN THE CALCIUM-DEPENDENT REGULATION OF
CC RHODOPSIN PHOSPHORYLATION. BINDS THREE CALCIUM IONS (BY
CC similarity).
CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BELONGS TO
CC THE RECOVERIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U91630; AAB51149.1; -.
DR EMBL; AF272896; AAG09045.1; -.
DR HSSP; P36610; 1G81.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001125; Recoverin.
DR Pfam; PF00036; ehand; 3
DR PRINTS; PR00450; RECOVERIN.
DR ProDom; PD00012; EF-hand; 1.
DR SMART; SM00054; EFh; 3.
DR PROSITE; PS00018; EF_HAND; 3.
KW Calcium-binding; Repeat; Myristate; Multigene family.
FT INIT_MET 0
FT BY SIMILARITY.
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
FT DOMAIN 35 46 ANCESTRAL CALCIUM SITE 1 (POTENTIAL).
FT CA_BIND 72 83 EF-HAND 2 (POTENTIAL).
FT CA_BIND 108 119 EF-HAND 3 (POTENTIAL).
FT CA_BIND 156 167 EF-HAND 4 (POTENTIAL).
SQ SEQUENCE 192 AA; 22130 MW; 807254352EC57003 CRC64;
Query Match 53.0%; Score 35; DB 1; Length 192;
Best Local Similarity 46.2%; Pred. No. 12;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 2 QLSLFXFTKFDLD 14
DB 98 QKLKWAFSMYDLD 110
| : : : :
| : : : :
RESULT 10
NCAD HUMAN
ID NCAD HUMAN STANDARD; PRT; 192 AA.
AC P29554; Q9H0W2;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

```

```

DE Neurocalcin delta.
GN NCALD.
OS Homo sapiens (Human), and
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606, 9913;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Human; TISSUE=Fetal brain;
RA MEDLINE=21167385; PubMed=11267673;
RA Wang W., Zhou Z., Zhao W., Huang Y., Tang R., Ying K., Xie Y., Mao Y.;
RT "Molecular cloning, mapping and characterization of the human
RT neurocalcin delta gene (NCALD).";
RL Biochim. Biophys. Acta 1518:162-167(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Human; TISSUE=Brain;
RA MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansoerge W., Boecker M., Bloecher H., Bauersachs S., Blum H.,
RA Lauber J., Dueterhoeft A., Beyer A., Koehrer K., Strack N.,
RA Meves H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs.";
RL Genome Res. 11:422-435(2001).
RN [3]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 106-111 AND 120-125.
RC SPECIES=Bovine; TISSUE=Brain;
RA MEDLINE=92287085; PubMed=1599450;
RA Okazaki K., Watanabe M., Ando Y., Hagiwara M., Terasawa M., Hidaka H.;
RT "Full sequence of neurocalcin, a novel calcium-binding protein
RT abundant in central nervous system.";
RL Biochem. Biophys. Res. Commun. 185:147-153(1992).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=Bovine; TISSUE=Brain;
RA MEDLINE=93247712; PubMed=8387172;
RA Hidaka H., Okazaki K.;
RT "Neurocalcin family: a novel calcium-binding protein abundant in
RT bovine central nervous system.";
RL Neurosci. Res. 16:73-77(1993).
RN [5]
RP CALCIUM-BINDING DATA.
RC SPECIES=Bovine;
RA Ladant D.;
RT "Calcium and membrane binding properties of bovine neurocalcin delta
RT expressed in Escherichia coli.";
RL J. Biol. Chem. 270:3179-3185(1995).
RN [6]
RP CRYSTALLIZATION.
RC SPECIES=Bovine;
RA MEDLINE=96407688; PubMed=8811741;
RA Kumar V.D., Hidaka H., Okazaki K., Vijay-Kumar S.;
RT "Crystallization and preliminary X-ray crystallographic studies of
RT recombinant bovine neurocalcin delta.";
RL Proteins 25:261-264(1996).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RC SPECIES=Bovine;
RA MEDLINE=99101387; PubMed=9886296;
RA Vijay-Kumar S., Kumar V.D.;
RT "Crystal structure of recombinant bovine neurocalcin.";
RL Nat. Struct. Biol. 6:80-88(1999).
CC -!- FUNCTION: MAY BE INVOLVED IN THE CALCIUM-DEPENDENT REGULATION OF
CC RHODOPSIN PHOSPHORYLATION. BINDS THREE CALCIUM IONS.
CC -!- TISSUE SPECIFICITY: RETINA, CEREBRUM, CEREBELLUM, BRAIN STEM,
CC SPINAL CORD, TESTIS, OVARY AND SMALL INTESTINE.
CC -!- MISCELLANEOUS: FIVE ISOPROTEIN FORMS OF NEUROCALCIN ARE DESIGNATED
CC ALPHA, BETA, GAMMA1, GAMMA2, AND DELTA.
CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BELONGS TO

```

```

CC THE RECOVERIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF251061; AAK34951.1; -
CC EMBL: AL136612; CAB66547.1; -
CC EMBL: D10884; BAA01706.1; -
CC PIR: JH0616; JH0616.
CC PDB: 1BJF; 22-JUL-99.
CC Genew: HGNC:7655; NCALD.
CC MIM: 606722; -
CC InterPro: IPR002048; EF-hand.
CC InterPro: IPR001125; Recoverin.
CC Pfam: PF00036; ehand; 3.
CC PRINTS: PR00450; RECOVERIN.
CC ProDom: PD000012; EF-hand; 1.
CC SMART: SM00054; EFh; 3.
CC PROSITE: PS00018; EF_HAND; 3.
CC Calcium-binding; Repeat; Myristate; Multigene family.
CC INIT_MET 0 0 BY SIMILARITY.
CC LIPID 1 1 MYRISTATE.
CC DOMAIN 35 46 ANCESTRAL CALCIUM SITE 1 (POTENTIAL).
CC CA_BIND 72 83 EF-HAND 2 (POTENTIAL).
CC CA_BIND 108 119 EF-HAND 3 (POTENTIAL).
CC CA_BIND 156 167 EF-HAND 4 (POTENTIAL).
CC CONFLICT 122 122 E -> V (IN REF. 2).
CC SEQUENCE 192 AA; 22114 MW; 807242F9F4D3BCD9 CRC64;
CC -----
CC Query Match 53.0%; Score 35; DB 1; Length 192;
CC Best Local Similarity 46.2%; Pred. No. 12;
CC Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
CC -----
CC QY 2 QSLSPXFTKFDLD 14
CC Db 98 OKLWAFSMTYDLD 110
CC -----
CC RESULT 11
CC NCAD_MOUSE STANDARD; PRT; 192 AA.
CC AC Q91X97;
CC DT 15-JUN-2002 (Rel. 41, Created)
CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE Neurocalcin delta.
CC GN NCALD.
CC OS Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC OX NCBI_TaxID=10090;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Salivary gland;
CC RA Straussberg R.;
CC RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC CC -!- FUNCTION: MAY BE INVOLVED IN THE CALCIUM-DEPENDENT REGULATION OF
CC CC RHODOPSIN PHOSPHORYLATION. BINDS THREE CALCIUM IONS (By
CC CC similarity).
CC CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BELONGS TO
CC CC THE RECOVERIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: BC011162; AAH11162.1; -
CC InterPro: IPR002048; EF-hand.
CC Pfam: PF00036; ehand; 3.
CC ProDom: PD000012; EF-hand; 1.
CC PROSITE: PS00018; EF_HAND; 3.
CC Calcium-binding; Repeat; Myristate; Multigene family.
CC INIT_MET 0 0 BY SIMILARITY.
CC LIPID 1 1 MYRISTATE (BY SIMILARITY).
CC DOMAIN 35 46 ANCESTRAL CALCIUM SITE 1 (POTENTIAL).
CC CA_BIND 72 83 EF-HAND 2 (POTENTIAL).
CC CA_BIND 108 119 EF-HAND 3 (POTENTIAL).
CC CA_BIND 156 167 EF-HAND 4 (POTENTIAL).
CC SEQUENCE 192 AA; 22100 MW; 807246EDA592BCD9 CRC64;
CC -----
CC Query Match 53.0%; Score 35; DB 1; Length 192;
CC Best Local Similarity 46.2%; Pred. No. 12;
CC Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
CC -----
CC QY 2 QSLSPXFTKFDLD 14
CC Db 98 OKLWAFSMTYDLD 110
CC -----
CC RESULT 12
CC VIS3_CHICK STANDARD; PRT; 192 AA.
CC AC P42324;
CC DT 01-NOV-1995 (Rel. 32, Created)
CC DT 01-NOV-1995 (Rel. 32, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE Rem-1 protein (Hippocalcin-like protein 1).
CC GN HPCAL1 OR REM1.
CC OS Gallus gallus (Chicken).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
CC OC Gallus.
CC OX NCBI_TaxID=9031;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=White leghorn;
CC RX MEDLINE=95215060; PubMed=7700627;
CC RA Kraut N., Frampton J., Graf T.;
CC RT "Rem-1, a putative direct target gene of the Myb-Ets fusion
CC RT oncoprotein in haematopoietic progenitors, is a member of the
CC RT recoverin family.";
CC RT Oncogene 10:1027-1036(1995).
CC RL -!- TISSUE SPECIFICITY: EXPRESSED IN HEMOPOIETIC CELLS, BUT ALSO IN
CC CC THE FIBROBLASTS, BONE MARROW, BRAIN, EYE AND GUT.
CC CC -!- INDUCTION: BY THE MYB-ETS FUSION ONCOGENE.
CC CC -!- MISCELLANEOUS: PROBABLY BINDS TWO OR THREE CALCIUM IONS (BY
CC CC SIMILARITY).
CC CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BELONGS TO
CC CC THE RECOVERIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X80875; CAA56843.1; -
CC HSP; P36610; IG81.
CC InterPro: IPR002048; EF-hand.
CC InterPro: IPR001125; Recoverin.
CC Pfam: PF00036; ehand; 3.
CC PRINTS: PR00450; RECOVERIN.
CC ProDom: PD000012; EF-hand; 1.
CC SMART: SM00054; EFh; 3.
CC PROSITE: PS00018; EF_HAND; 3.
CC -----

```

KW Calcium-binding; Repeat; Myristate.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT LIPID 1 1 MYRISTATE (BY SIMILARITY).  
 FT DOMAIN 35 46 ANCESTRAL CALCIUM SITE 1 (POTENTIAL).  
 FT CA\_BIND 72 83 EF-HAND 2 (POTENTIAL).  
 FT CA\_BIND 108 119 EF-HAND 3 (POTENTIAL).  
 FT CA\_BIND 156 167 EF-HAND 4 (POTENTIAL).  
 SQ SEQUENCE 192 AA; BE7221EA537332FA CRC64;  
  
 Query Match 53.0%; Score 35; DB 1; Length 192;  
 Best Local Similarity 46.2%; Pred. No. 12;  
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
  
 QY 2 QSLSPXFTKFDLD 14  
 | | | | : | | |  
 Db 98 QKLKWAFFSMYDLD 110  
  
 RESULT 13  
 VIS3\_HUMAN STANDARD; PRT; 192 AA.  
 AC P37235; Q969S5;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE Visinin-like protein 3 (VILIP-3) (Calcium-binding protein BDR-1)  
 DE (HLP2) (Hippocalcin-like protein 1).  
 GN HPCALL1 OR BDR1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=94312451; PubMed=8038222;  
 RA Kobayashi M., Takamatsu K., Fujishiro M., Saitoh S., Noguchi T.;  
 RT "Molecular cloning of a novel calcium-binding protein structurally  
 related to hippocalcin from human brain and chromosomal mapping of  
 its gene.";  
 RL Biochim. Biophys. Acta 1222:515-518(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain, and Pancreas;  
 RA Strausberg R.;  
 RT Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: MAY BE INVOLVED IN THE CALCIUM-DEPENDENT REGULATION OF  
 RHODOPSIN PHOSPHORYLATION.  
 CC -!- MISCELLANEOUS: PROBABLY BINDS TWO OR THREE CALCIUM IONS (BY  
 SIMILARITY).  
 CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BELONGS TO  
 THE RECOVERIN SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; D16227; BAA03754.1;  
 DR EMBL; BC009846; AAH09846.1;  
 DR EMBL; BC017028; AAH17028.1;  
 DR EMBL; BC017482; AAH17482.1;  
 DR HSP; P36610; 1G8I.  
 DR Genew; HGNC:5145; HPCALL1.  
 DR MIM; 600207;  
 DR InterPro; IPR002048; EF-hand.  
 DR InterPro; IPR001125; Recoverin.  
 DR Pfam; PF00036; ehand; 3.  
 DR PRINTS; PR00450; RECOVERIN.  
 DR ProDom; PD000012; EF-hand; 1.

DR SMART; SM00054; EFh; 3.  
 DR PROSITE; PS00018; EF\_HAND; 3.  
 KW Calcium-binding; Repeat; Myristate.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT LIPID 1 1 MYRISTATE (POTENTIAL).  
 FT DOMAIN 35 46 ANCESTRAL CALCIUM SITE 1.  
 FT CA\_BIND 72 83 EF-HAND 2 (POTENTIAL).  
 FT CA\_BIND 108 119 EF-HAND 3 (POTENTIAL).  
 FT CA\_BIND 156 167 EF-HAND 4 (POTENTIAL).  
 FT CONFLICT 18 18 N -> K (IN REF. 1).  
 FT CONFLICT 87 87 N -> G (IN REF. 1).  
 FT CONFLICT 92 93 SR -> RG (IN REF. 1).  
 SQ SEQUENCE 192 AA; 22182 MW; 277320AEDC02AE79 CRC64;  
  
 Query Match 53.0%; Score 35; DB 1; Length 192;  
 Best Local Similarity 46.2%; Pred. No. 12;  
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
  
 QY 2 QSLSPXFTKFDLD 14  
 | | | | : | | |  
 Db 98 QKLKWAFFSMYDLD 110  
  
 RESULT 14  
 VIS3\_MOUSE STANDARD; PRT; 192 AA.  
 AC P35333;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Visinin-like protein 3 (VILIP-3) (Neural visinin-like protein 3)  
 DE (NVL-3) (NVP-3) (Hippocalcin-like protein 1).  
 GN HPCALL1.  
 OS Mus musculus (Mouse), and  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090, 10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Mouse; STRAIN=BALE/c; TISSUE=Brain;  
 RA Ritter B., Modregger J., Plomann M.;  
 RT "Interactions of the murine neural visinin-like protein 3 (mNVP-3).";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Mouse; TISSUE=Breast tumor;  
 RA Strausberg R.;  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Rat; TISSUE=Brain;  
 RX MEDLINE=93367470; PubMed=8360675;  
 RA Kajimoto Y., Shirai Y., Mukai H., Kuno T., Tanaka C.;  
 RT "Molecular cloning of two additional members of the neural visinin-  
 like Ca(2+)-binding protein gene family.";  
 RL J. Neurochem. 61:1091-1096(1993).  
 CC -!- FUNCTION: MAY BE INVOLVED IN THE CALCIUM-DEPENDENT REGULATION OF  
 RHODOPSIN PHOSPHORYLATION.  
 CC -!- TISSUE SPECIFICITY: IN NEURONAL CELLS, BUT NOT AS SPECIFICALLY AS  
 VILIP-1 OR VILIP-2.  
 CC -!- MISCELLANEOUS: PROBABLY BINDS TWO OR THREE CALCIUM IONS (BY  
 SIMILARITY).  
 CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BELONGS TO  
 THE RECOVERIN SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC EMBL; AF085192; AAC35552.1; -
DR EMBL; BC001997; AA01097.1; -
DR EMBL; D13126; BAA02428.1; -
DR PIR; JH0816; JH0816.
DR HSSP; P36610; IG81.
DR MGD; MGI:1855689; Hpcall.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF000138; lectin_legA.
DR PRINTS; PR00450; RECOVERIN.
DR PRODOM; PD000012; EF-hand; 1.
DR SMART; SM00054; EFH; 3.
DR PROSITE; PS00018; EF-HAND; 3.
KW Calcium-binding; Repeat; Myristate.
FT INIT_MET 0 BY SIMILARITY.
FT LIPID 1 MYRISTATE (BY SIMILARITY).
FT DOMAIN 35 46 ANCESTRAL CALCIUM SITE 1 (POTENTIAL).
FT CA_BIND 72 83 EF-HAND 2 (POTENTIAL).
FT CA_BIND 108 119 EF-HAND 3 (POTENTIAL).
FT CA_BIND 156 167 EF-HAND 4 (POTENTIAL).
SQ SEQUENCE 192 AA; 22207 MW; 178ADDD3F89B097B CRC64;

```

Query Match 53.0%; Score 35; DB 1; Length 192;  
 Best Local Similarity 46.2%; Pred. No. 12;  
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```

QY 2 QSLSFXTKFDLD 14
   | | | | |
DB 98 OKLWAFSMYDLD 110

```

```

RESULT 15
LECA_CRAFL STANDARD; PRT; 236 AA.
AC P81517; P81636;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lentin alpha chain [Contains: Lentin beta chain; Lentin gamma chain].
OS Cratylia floribunda.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Cratylia.
OX NCBI_TaxID=83131;
RN [1]
RP TISSUE=Seed;
RC Cavada B.S., Nogueira N.A.P., Farias C.M.A.S., Grangeiro T.B.,
RA Romas M.V., Thole H.H., Raida M., Rouge P., Calvete J.J.;
RT "Primary structure and kinetic interaction with glycoproteins of the
RT lectin from seeds of Cratylia floribunda.";
RL Protein Pept. Lett. 6:27-34(1999).
RN [2]
RP SEQUENCE, MASS SPECTROMETRY, AND X-RAY CRYSTALLOGRAPHY (3.5
RP ANGSTROMS).
RC TISSUE=Seed;
RX MEDLINE=99184792; PubMed=10082964;
RA Calvete J.J., Thole H.H., Raida M., Urbanke C., Romero A.,
RA Grangeiro T.B., Ramos M.V., Almeida da Rocha I.M., Guimaraes F.N.,
RA Cavada B.S.;
RT "Molecular characterization and crystallization of Diocleinae
RT lectins.";
RL Biochim. Biophys. Acta 1430:367-375(1999).
CC -!- FUNCTION: D-Mannose/D-Glucose-binding lectin. Mixture of 60% alpha
CC lectin and 40% of its beta and gamma proteolytic fragments.
CC -!- COFACTOR: Requires calcium and manganese ions for full activity.
CC -!- SUBUNIT: PH-DEPENDENT HOMODIMER OF ALPHA CHAINS AT PH VALUES LOWER
CC THAN 6.5 AND A HOMOTETRAMER OF ALPHA CHAINS ABOVE THIS VALUE.
CC -!- TISSUE SPECIFICITY: SEED.
CC -!- PTM: The beta and gamma chains are produced by partial proteolytic
CC processing of the lectin alpha chain by an asparaginyl
CC endopeptidase.

```

```

CC -!- MASS SPECTROMETRY: MW=25397; MW_ERR=3; METHOD=Electrospray;
CC RANGE=1-236.
CC -!- MASS SPECTROMETRY: MW=12847; MW_ERR=2; METHOD=Electrospray;
CC RANGE=1-118.
CC -!- MASS SPECTROMETRY: MW=12568; MW_ERR=2; METHOD=Electrospray;
CC RANGE=119-236
CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
CC HSSP; P02866; LDQ2.
CC InterPro; IPR000985; Lentin_legA.
CC InterPro; IPR001220; Lentin_legB.
CC Pfam; PF00138; lectin_legA; 1.
CC Pfam; PF00139; lectin_legB; 2.
CC ProDom; PD000671; Lentin_legA; 1.
CC ProDom; PD000711; Lentin_legB; 2.
CC PROSITE; PS00307; LECTIN_LEGUME_BETA; 1..
CC PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1..
KW Lentin; Calcium; Manganese.
FT CHAIN 1 236 LECTIN ALPHA CHAIN.
FT CHAIN 1 118 LECTIN BETA CHAIN.
FT CHAIN 119 236 LECTIN GAMMA CHAIN.
SQ SEQUENCE 236 AA; 25398 MW; CA73BC77F4324251 CRC64;

```

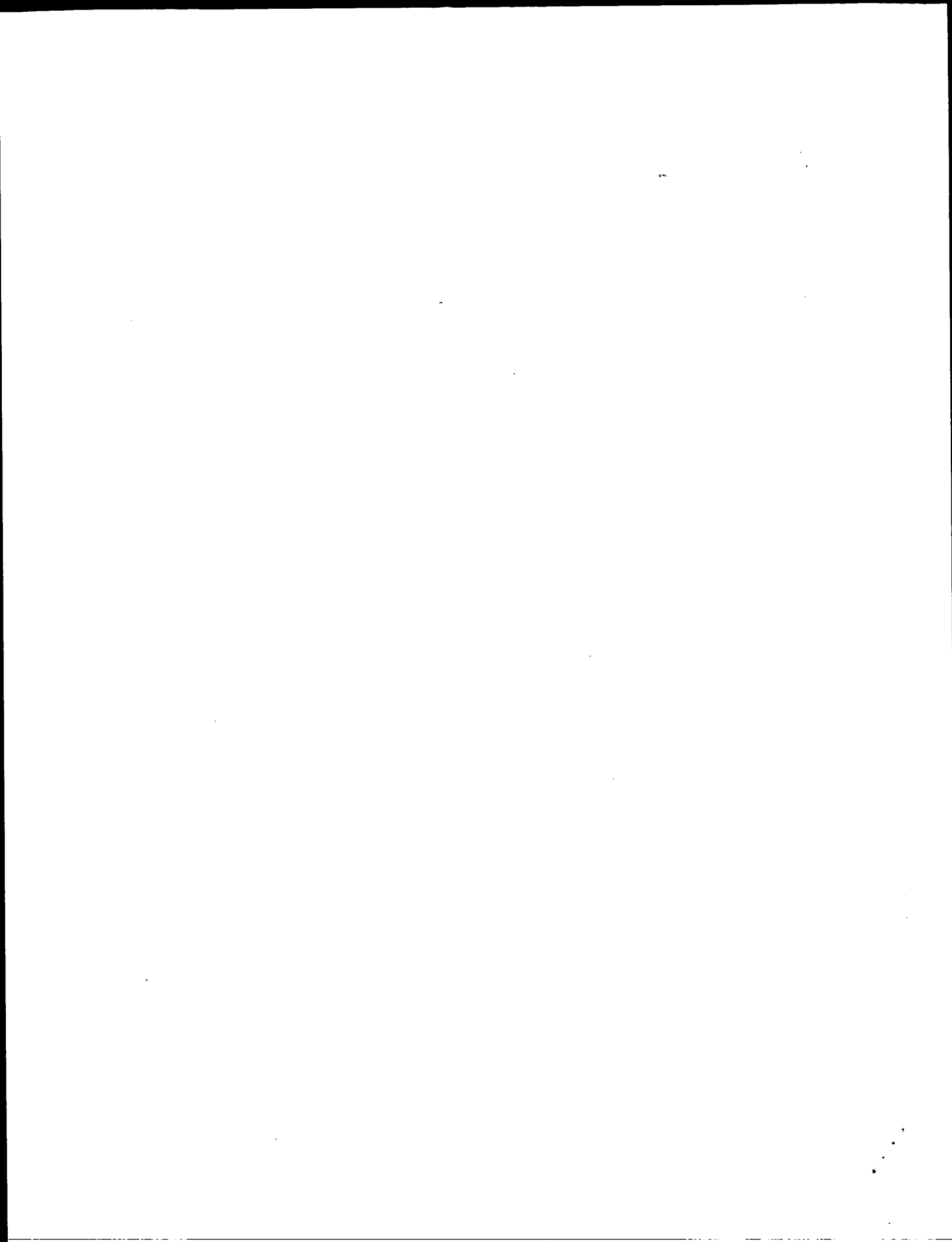
Query Match 53.0%; Score 35; DB 1; Length 236;  
 Best Local Similarity 63.6%; Pred. No. 15;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 AQSLSFXFTKF 11
   | | | | |
DB 123 AQSLSHFTNQF 133

```

Search completed: February 26, 2003, 14:49:00  
 Job time : 12 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:48:41 ; Search time 29 Seconds  
(without alignments)

Title: US-09-476-485A-31  
Perfect score: 66  
Sequence: 1 AOSLSFXFTKFDLD 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

```

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45

```

Database : SPTREMBL 21:\*

- ```

1:  sp:archaea.*
2:  sp:archaea.*
3:  sp:bacteria.*
4:  sp:fungi.*
5:  sp:human.*
6:  sp:invertebrate.*
7:  sp:mammal.*
8:  sp:mhc.*
9:  sp:organelle.*
10: sp:phase.*
11: sp:plant.*
12: sp:rodent.*
13: sp:virus.*
14: sp:vertebrate.*
15: sp:unclassified.*
16: sp_rvrius.*
17: sp_bacteriap.*
18: sp_archaeap.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query % |          | Length | DB | ID     | Description        |
|------------|-------|---------|----------|--------|----|--------|--------------------|
|            |       | Match   | Mismatch |        |    |        |                    |
| 1          | 64    | 97.0    | 0        | 279    | 10 | Q9M7M4 | phaseolus v        |
| 2          | 55    | 83.3    | 3        | 272    | 10 | Q9ZTA9 | Q9Zta9 dolichos la |
| 3          | 42    | 63.6    | 4        | 408    | 16 | Q91LY8 | Q91ly8 pseudomonas |
| 4          | 42    | 63.6    | 4        | 896    | 5  | Q9VVV9 | Q9vvv9 drosophila  |
| 5          | 42    | 63.6    | 4        | 928    | 5  | Q9N1H2 | Q9nih2 drosophila  |
| 6          | 40    | 60.6    | 6        | 284    | 10 | Q9FYU9 | Q9fyu9 sophora fla |
| 7          | 40    | 60.6    | 6        | 292    | 12 | Q65239 | Q65239 african swi |
| 8          | 40    | 60.6    | 6        | 301    | 12 | Q65196 | Q65196 african swi |
| 9          | 39    | 59.1    | 7        | 279    | 10 | Q9ZWP4 | Q9zwp4 robinia pse |
| 10         | 39    | 59.1    | 7        | 285    | 10 | Q9ZWP5 | Q9zwp5 robinia pse |
| 11         | 39    | 59.1    | 7        | 785    | 5  | Q962M0 | Q962m0 plasmodium  |
| 12         | 39    | 59.1    | 7        | 924    | 16 | Q91SU2 | Q91su2 pseudomonas |
| 13         | 38    | 57.6    | 8        | 254    | 10 | Q43376 | Q43376 arachis hyp |
| 14         | 38    | 57.6    | 8        | 467    | 5  | Q20265 | Q20265 caenorhabdi |
| 15         | 38    | 57.6    | 8        | 705    | 5  | Q23024 | Q23024 caenorhabdi |
| 16         | 38    | 57.6    | 8        | 708    | 5  | Q93380 | Q93380 caenorhabdi |

|    |    |      |      |    |        |                     |
|----|----|------|------|----|--------|---------------------|
| 17 | 38 | 57.6 | 709  | 10 | Q9FHY2 | Q9fhy2 arabidopsis  |
| 18 | 38 | 57.6 | 807  | 5  | Q18514 | Q18514 caenorhabdi  |
| 19 | 38 | 57.6 | 2656 | 5  | Q9GNU3 | Q9gnu3 paracentrot  |
| 20 | 37 | 56.1 | 99   | 2  | Q8RM68 | Q8rm68 bacteroides  |
| 21 | 37 | 56.1 | 220  | 3  | O42911 | O42911 schizosacch  |
| 22 | 37 | 56.1 | 298  | 16 | P73352 | P73352 synecocyst   |
| 23 | 37 | 56.1 | 346  | 5  | Q9NW1  | Q9nw1 caenorhabdi   |
| 24 | 37 | 56.1 | 462  | 5  | Q9TVM1 | Q9tvm1 toxoplasma   |
| 25 | 36 | 54.5 | 129  | 12 | O11290 | O11290 mollusum c   |
| 26 | 36 | 54.5 | 284  | 5  | O17957 | O17957 caenorhabdi  |
| 27 | 36 | 54.5 | 318  | 16 | Q8RDV1 | Q8rdv1 fusobacteri  |
| 28 | 36 | 54.5 | 425  | 16 | Q9A454 | Q9a454 caulobacter  |
| 29 | 36 | 54.5 | 450  | 2  | Q9E292 | Q9e292 lactococcus  |
| 30 | 36 | 54.5 | 534  | 12 | Q38195 | Q38195 mollusum c   |
| 31 | 36 | 54.5 | 590  | 16 | Q9RWT8 | Q9rwt8 leishmanias  |
| 32 | 36 | 54.5 | 699  | 5  | O76608 | O76608 caenorhabdi  |
| 33 | 36 | 54.5 | 759  | 5  | Q95P54 | Q95p54 lymanaea sta |
| 34 | 35 | 53.8 | 565  | 8  | Q9P7K4 | Q9p7k4 crassostrea  |
| 35 | 35 | 53.0 | 25   | 10 | Q9S8C1 | Q9s8c1 griffonia s  |
| 36 | 35 | 53.0 | 90   | 6  | Q9SLH3 | Q9slh3 mustela put  |
| 37 | 35 | 53.0 | 110  | 4  | O75544 | O75544 homo sapien  |
| 38 | 35 | 53.0 | 124  | 3  | Q07438 | Q07438 saccharomyc  |
| 39 | 35 | 53.0 | 170  | 4  | O8WYJ4 | O8wyj4 homo sapien  |
| 40 | 35 | 53.0 | 173  | 4  | O86L23 | Q86l23 homo sapien  |
| 41 | 35 | 53.0 | 189  | 5  | Q9VNF9 | Q9vnf9 drosophila   |
| 42 | 35 | 53.0 | 191  | 6  | Q95KJ8 | Q95kj8 macaca fasc  |
| 43 | 35 | 53.0 | 232  | 10 | Q94719 | Q94719 triticum tu  |
| 44 | 35 | 53.0 | 266  | 10 | O8RVX4 | O8rvx4 phaseolus v  |
| 45 | 35 | 53.0 | 270  | 10 | O40987 | O40987 pisum sativ  |

## ALIGNMENTS

## RESULT 1

|                                                                        |                                         |                    |                                   |
|------------------------------------------------------------------------|-----------------------------------------|--------------------|-----------------------------------|
|                                                                        | PRT;                                    | 279                | AA.                               |
| PRELIMINARY;                                                           |                                         |                    |                                   |
| Q9W7M4                                                                 |                                         |                    |                                   |
| ID                                                                     |                                         |                    |                                   |
| AC                                                                     |                                         |                    |                                   |
| Q9W7M4;                                                                |                                         |                    |                                   |
| DT                                                                     |                                         |                    |                                   |
| 01-OCT-2000                                                            | (TREMBLrel. 15, Created)                |                    |                                   |
| DT                                                                     |                                         |                    |                                   |
| 01-OCT-2000                                                            | (TREMBLrel. 15, Last sequence update)   |                    |                                   |
| DT                                                                     |                                         |                    |                                   |
| 01-MAR-2002                                                            | (TREMBLrel. 20, Last annotation update) |                    |                                   |
| DT                                                                     |                                         |                    |                                   |
| Mannose lectin FRIL {Fragment}.                                        |                                         |                    |                                   |
| OOS                                                                    |                                         |                    |                                   |
| Phaseolus vulgaris (Kidney bean) (French bean).                        |                                         |                    |                                   |
| OC                                                                     |                                         |                    |                                   |
| Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;     |                                         |                    |                                   |
| OC                                                                     |                                         |                    |                                   |
| Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  |                                         |                    |                                   |
| OC                                                                     |                                         |                    |                                   |
| eutrosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus. |                                         |                    |                                   |
| NCBI_TaxID=3885;                                                       |                                         |                    |                                   |
| [1]                                                                    |                                         |                    |                                   |
| SEQUENCE FROM N.A.                                                     |                                         |                    |                                   |
| RNP                                                                    |                                         |                    |                                   |
| MOORE J.G.; COLUCCI G.; FUCHS C.A.; HICKLIN D.J., CHRISPEELS M.J.,     |                                         |                    |                                   |
| FELDMAN M.;                                                            |                                         |                    |                                   |
| "A new lectin in red kidney bean called PVFRIL stimulates              |                                         |                    |                                   |
| proliferation of NIH3T3 cells expressing the Flt3 receptor."           |                                         |                    |                                   |
| Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.               |                                         |                    |                                   |
| EMBL; AF121459; AAF28739.1; -;                                         |                                         |                    |                                   |
| HSSP; P02866; 1ONA.                                                    |                                         |                    |                                   |
| InterPro; IPRO00985; Lectin_legA.                                      |                                         |                    |                                   |
| InterPro; IPRO01220; Lectin_legB.                                      |                                         |                    |                                   |
| Pfam; PF00138; lectin_legA; 1.                                         |                                         |                    |                                   |
| Pfam; PF00139; lectin_legB; 1.                                         |                                         |                    |                                   |
| ProDom; PD000671; Lectin_legA; 1.                                      |                                         |                    |                                   |
| ProDom; PD000711; Lectin_legB; 1.                                      |                                         |                    |                                   |
| PROSITE; PS00307; LECTIN_LEGUME_ALPHA; 1.                              |                                         |                    |                                   |
| PROSITE; PS00308; LECTIN_LEGUME_BETA; UNKNOWN_1.                       |                                         |                    |                                   |
| Lectin.                                                                |                                         |                    |                                   |
| NON_TER                                                                |                                         |                    |                                   |
| SEQUENCE                                                               | 279                                     | AA;                | 31102 MW; F8919CF8B3EE4652 CRC64; |
|                                                                        | 1                                       |                    | 1                                 |
| Query Match                                                            | 97.0%                                   | Score 64;          | DB 10; Length 279;                |
| Best Local Similarity                                                  | 92.9%;                                  | Pred. No. 0.00015; |                                   |
| Matches 13;                                                            | Conservative                            | 0;                 | Mismatches 1; Indels 0; Gaps 0;   |
| I                                                                      |                                         |                    |                                   |
| Q9W7M4                                                                 |                                         |                    |                                   |
| AQSLSFXTKFDLD                                                          | 14                                      |                    |                                   |

```

||||||| |||||||
1 AQSLSFNTKFDLD 14

RESULT 2
Q9ZTA9          PRELIMINARY;      PRT;      272 AA.
AC Q9ZTA9;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Mannoese lectin.
GN FRIL.
OS Dolichos lab lab (field bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Dolichos.
OX NCBI_TaxID=35936;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COTYLEDON;
RC MEDLINE=99110944; PubMed=9892587;
RA Colucci G., Moore J.G., Feldman M., Chrispeels M.J.;
RT "cDNA cloning of FRIL, a lectin from Dolichos lablab, that preserves
RT hemagglutinating activity in suspension culture.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:646-650(1999).
DR EMBL: AF067417; AAD10734.1;
DR HSSP: P02866; IONA.
DR InterPro: IPR000985; Lectin_legA.
DR InterPro: IPR001220; Lectin_legB.
DR Pfam: PF00138; lectin_legA; 1.
DR Pfam: PF00139; lectin_legB; 1.
DR ProDom: PD000671; Lectin_legA; 1.
DR ProDom: PD000711; Lectin_legB; 1.
DR PROSITE: PS00308; LECTIN_LEGUME_ALPHA; 1.
DR PROSITE: PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
RW Lectin.
SQ SEQUENCE 272 AA; 29900 MW; EA6C004307441495 CRC64;

Query Match 83.3%; Score 55; DB 10; Length 272;
Best Local Similarity 91.7%; Pred. No. 0.0086;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXTKFD 12
   ||||| |||||
DB 9 AQSLSFSTKFD 20

RESULT 3
Q9IIY8          PRELIMINARY;      PRT;      408 AA.
AC Q9IIY8;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE Hypothetical protein PA2127.
GN PA2127.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Barber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";

```

```

RL Nature 406:959-964(2000).
DR EMBL: AE004640; AAG05515.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 408 AA; 47591 MW; 87FD5B3B704643B1 CRC64;

Query Match 63.6%; Score 42; DB 16; Length 408;
Best Local Similarity 57.1%; Pred. No. 4.8;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSFXTKFDLD 14
   ||||| |||||
DB 385 ARSLSFQSOADID 398

RESULT 4
Q9VVV9          PRELIMINARY;      PRT;      896 AA.
AC Q9VVV9; Q9VVW0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Nkd protein.
GN Nkd OR CG11614 OR CG18224.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Puskas J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Stadling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003518; AAF49198.2;
DR HSSP: P36610; 1G81.
DR FlyBase: FBgn0002945; nkd.

```



SQ SEQUENCE 896 AA; 98892 MW; 3DC2AB6F07D246BB CRC64;

Query Match 63.6%; Score 42; DB 5; Length 896;  
Best Local Similarity 57.1%; Pred. No. 11;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFDLD 14  
: | | | | | | | | | |  
Db 158 SQPLQFSFTFYDLD 171

## RESULT 5

Q9NIH2 PRELIMINARY; PRT; 928 AA.  
ID Q9NIH2  
AC Q9NIH2  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Naked cuticle.  
GN NKD OR CG11614 OR CG18224.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Prorygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zeng W., Wharton K.A. Jr., Mack J.A., Wang K., Gadbaw M., Suyama K., Klein P.S., Scott M.P.;  
RA "Naked cuticle encodes an inducible antagonist of Wnt signalling.";  
RL Nature 0:0-0(2000).  
DR EMBL; AF213376; AAF34825.1; -.  
DR HSSP; P36610; IG81.  
DR FlyBase; FBgn0002945; nkd.  
SQ SEQUENCE 928 AA; 102591 MW; D543F378E2BC5122 CRC64;

Query Match 63.6%; Score 42; DB 5; Length 928;  
Best Local Similarity 57.1%; Pred. No. 11;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFDLD 14  
: | | | | | | | | | |  
Db 190 SQPLQFSFTFYDLD 203

## RESULT 6

Q9FYU9 PRELIMINARY; PRT; 284 AA.  
ID Q9FYU9  
AC Q9FYU9  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Lectin.  
OS Sophora flavescens.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Sophora.  
OX NCBI\_TaxID=49840;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=ROOT;  
RA Yan B., Ma Z.G., Wang L.X., Chai H.M., Huang X.Q.;  
RT "Cloning and sequencing of Sophora flavescens lectin gene, 284 AA.";  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF285121; AAG00508.1; -.  
DR HSSP; P02866; 1DQ2.  
DR InterPro; IPR000985; Lectin\_legA.  
DR InterPro; IPR001220; Lectin\_legB.  
DR Pfam; PF00138; lectin\_legA; 1.  
DR Pfam; PF00139; lectin\_legB; 1.  
DR ProDom; PD000671; Lectin\_legA; 1.  
DR ProDom; PD000711; Lectin\_legB; 1.  
DR PROSITE; PS00308; LECTIN\_LEGUME\_ALPHA; 1.

DR PROSITE; PS00307; LECTIN\_LEGUME\_BETA; UNKNOWN; 1.  
SQ SEQUENCE 284 AA; 31318 MW; 2DC947EB3CBE0FB2 CRC64;

Query Match 60.6%; Score 40; DB 10; Length 284;  
Best Local Similarity 66.7%; Pred. No. 8.3;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFD 12  
: | | | | | | | | | |  
Db 31 ADSLSFTSFDFD 42

## RESULT 7

Q65239 PRELIMINARY; PRT; 292 AA.  
ID Q65239  
AC Q65239  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE ORF\_115R.  
OS African swine fever virus (ASFV).  
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.  
OX NCBI\_TaxID=10497;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MALAWI LIL20 /1;  
RX MEDLINE=94014996; PubMed=8409937;  
RA Vydelingum S., Baylis S.A., Bristow C., Smith G.L., Dixon L.K.;  
RT "Duplicated genes within the variable right end of the genome of a pathogenic isolate of African swine fever virus.";  
RL J. Gen. Virol. 74:2125-2130(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MALAWI LIL20 /1;  
RA Dixon L.K., Twigg S.R.F., Baylis S.A., Vydelingum S., Bristow C., Hammond J.M., Smith G.L.;  
RT "Nucleotide sequence of a 55 kbp region from the right end of the genome of a pathogenic African swine fever virus isolate (Malawi LIL20/1).";  
RL J. Gen. Virol. 7:1655-1684(1994).  
DR EMBL; X71982; CAA50835.1; -.  
SQ SEQUENCE 292 AA; 34549 MW; CA61A30E49D9219D CRC64;

Query Match 60.6%; Score 40; DB 12; Length 292;  
Best Local Similarity 53.8%; Pred. No. 8.6;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 QSLSFXTKFDLD 14  
: | | | | | | | | | |  
Db 135 ETLXFTFDID 147

## RESULT 8

Q65196 PRELIMINARY; PRT; 301 AA.  
ID Q65196  
AC Q65196  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Proliferating cell nuclear antigen-like protein.  
GN E30LR.  
OS African swine fever virus (ASFV).  
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.  
OX NCBI\_TaxID=10497;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BA7IV;  
RX MEDLINE=21820291; PubMed=11831707;  
RA Yanez R.J., Rodriguez J.M., Nogal M.L., Yuste L., Enriquez C., Rodriguez J.F., Vinuela E.;  
RT "Analysis of the complete nucleotide sequence of African swine fever virus.";  
RL Virology 208:249-278(1995).

[2]  
 RX SEQUENCE FROM N.A.  
 RA STRAIN=BA71V;  
 RC MEDLINE=94233765; PubMed=8178480;  
 RA De la Vega I., Gonzalez A., Blasco R., Calvo V., Vinuela E.;  
 RT "Nucleotide sequence and variability of the inverted terminal  
 RL repetitions of African swine fever virus DNA.";  
 RN Virology 201:152-156(1994).  
 [3]  
 RX SEQUENCE FROM N.A.  
 RA STRAIN=BA71V;  
 RC MEDLINE=90219205; PubMed=2325203;  
 RA Gonzalez A., Calvo V., Almazan F., Almendral J.M., Ramirez J.C.,  
 RA De La Vega I., Blasco R., Vinuela E.;  
 RT "Multigene families in African swine fever virus: Family 360.";  
 RL J. Virol. 64:2073-2081(1990).  
 [4]  
 RX SEQUENCE FROM N.A.  
 RA STRAIN=BA71V;  
 RC MEDLINE=90219204; PubMed=2325202;  
 RA Almendral J.M., Almazan F., Blasco R., Vinuela E.;  
 RT "Multigene families in African swine fever virus: Family 110.";  
 RL J. Virol. 64:2064-2072(1990).  
 [5]  
 RX SEQUENCE FROM N.A.  
 RA STRAIN=BA71V;  
 RC MEDLINE=91134988; PubMed=1994575;  
 RA Canacho A., Vinuela E.;  
 RT "Protein p22 of african swine fever virus: an early structural protein  
 RL that is incorporated into the membrane of infected cells.";  
 RN Virology 181:251-257(1991).  
 [6]  
 RX SEQUENCE FROM N.A.  
 RA STRAIN=BA71V;  
 RC STRAIN=BA71V;  
 RA Almazan F., Murguía J.R., Rodriguez J.M., La Vega I., Vinuela E.;  
 RT "A set of African swine fever virus tandem repeats shares similarities  
 RL with SAR-like sequences.";  
 RN J. Gen. Virol. 0:0-0(0).  
 [7]  
 RX SEQUENCE FROM N.A.  
 RA STRAIN=BA71V;  
 RC MEDLINE=94187118; PubMed=8139051;  
 RA Rodriguez J.M., Yanez R.J., Pan R., Rodriguez J.F., Salas M.L.,  
 RA Vinuela E.;  
 RT "Multigene families in African swine fever virus: family 505.";  
 RL J. Virol. 68:2746-2751(1994).  
 [8]  
 RX SEQUENCE FROM N.A.  
 RA STRAIN=BA71V;  
 RC MEDLINE=93346971; PubMed=8393914;  
 RA Yanez R.J., Rodriguez J.M., Rodriguez J.F., Salas M.L., Vinuela E.;  
 RT "African swine fever virus thymidylate kinase gene: sequence and  
 RL transcriptional mapping.";  
 RN J. Gen. Virol. 74:1633-1638(1993).  
 [9]  
 RX SEQUENCE FROM N.A.  
 RA STRAIN=BA71V;  
 RC MEDLINE=94065656; PubMed=8245848;  
 RA Alcamí A., Angulo A., Vinuela E.;  
 RT "Mapping and sequence of the gene encoding the African swine fever  
 RL virion protein of M(r) 11500.";  
 RN J. Gen. Virol. 74:2317-2324(1993).  
 [10]  
 RX SEQUENCE FROM N.A.  
 RA STRAIN=BA71V;  
 RC MEDLINE=93277388; PubMed=8503790;  
 RA Munoz M., Freije J.M., Salas M.L., Vinuela E., Lopez-Otin C.;  
 RT "Structure and expression in E. coli of the gene coding for protein  
 RL p10 of African swine fever virus.";  
 RN Arch. Virol. 130:93-107(1993).  
 [11]  
 RX SEQUENCE FROM N.A.  
 RA STRAIN=BA71V;  
 RC MEDLINE=90357780; PubMed=2389555;  
 RA Blasco R., Lopez-Otin C., Munoz M., Bockamp E.-O.O., Simon-Mateo C.,  
 RA Vinuela E.;  
 RT "Sequence and evolutionary relationships of African swine fever virus  
 RL thymidine kinase.";  
 RN Virology 178:301-304(1990).  
 [12]  
 RX SEQUENCE FROM N.A.  
 RA STRAIN=BA71V;  
 RC MEDLINE=93281390; PubMed=8506138;  
 RA Yanez R.J., Boursnell M., Nogal M.L., Yuste L., Vinuela E.;  
 RT "African swine fever virus encodes two genes which share significant  
 RL homology with the two largest subunits of DNA-dependent RNA  
 RN polymerases.";  
 RN Nucleic Acids Res. 21:2423-2427(1993).  
 [13]  
 RX SEQUENCE FROM N.A.  
 RA STRAIN=BA71V;  
 RC MEDLINE=93353606; PubMed=8102411;  
 RA Yanez R.J., Almazan F., Vinuela E., Rodriguez J.F.;  
 RT "African swine fever virus encodes a CD2 homolog responsible for the  
 RL adhesion of erythrocytes to infected cells.";  
 RN J. Virol. 67:5312-5320(1993).  
 [14]  
 RX SEQUENCE FROM N.A.  
 RA STRAIN=BA71V;  
 RC MEDLINE=94085774; PubMed=8262374;  
 RA Yanez R.J., Rodriguez J.M., Boursnell M.E.G., Rodriguez J.F.,  
 RA Vinuela E.;  
 RT "Two putative african swine fever virus helicases similar to yeast  
 RL 'DEAH' pre-mRNA processing proteins and vaccinia virus atpases D1L and  
 RN D6R.";  
 RN Gene 134:161-174(1993).  
 [15]  
 RX SEQUENCE FROM N.A.  
 RA STRAIN=BA71V;  
 RC MEDLINE=90223993; PubMed=2327074;  
 RA Lopez-Otin C., Freije J.M.P., Parra F., Mendez E., Vinuela E.;  
 RT "Mapping and sequence of the gene coding for protein p72, the major  
 RL capsid protein of African swine fever virus.";  
 RN Virology 175:477-484(1990).  
 [16]  
 RX SEQUENCE FROM N.A.  
 RA STRAIN=BA71V;  
 RC MEDLINE=94123986; PubMed=8293992;  
 RA Rodriguez J.M., Yanez R.J., Rodriguez J.F., Vinuela E., Salas M.L.;  
 RT "The DNA polymerase-encoding gene of African swine fever virus:  
 RL sequence and transcriptional mapping.";  
 RN Gene 136:103-110(1993).  
 [17]  
 RX SEQUENCE FROM N.A.  
 RA STRAIN=BA71V;  
 RC MEDLINE=93327788; PubMed=8335009;  
 RA Simon-Mateo C., Andres G., Vinuela E.;  
 RT "Polyprotein processing in African swine fever virus: a novel gene  
 RL expression strategy for a DNA virus.";  
 RN EMBO J. 12:2977-2987(1993).  
 [18]  
 RX SEQUENCE FROM N.A.  
 RA STRAIN=BA71V;  
 RC MEDLINE=93233210; PubMed=8474154;  
 RA Prados F.J., Vinuela E., Alcamí A.;  
 RT "Sequence and characterization of the major early phosphoprotein p32  
 RL of African swine fever virus.";  
 RN J. Virol. 67:2475-2485(1993).  
 [19]  
 RX SEQUENCE FROM N.A.  
 RA STRAIN=BA71V;  
 RC MEDLINE=92260660; PubMed=1583732;  
 RA Alcamí A., Angulo A., Lopez-Otin C., Munoz M., Freije J.M.P.,  
 RA Carrascosa A.L., Vinuela E.;  
 RT "Amino acid sequence and structural properties of protein p12, an  
 RL African swine fever virus attachment protein.";



```

DR InterPro: IPR003959; AAA_ATPase_centr.
DR InterPro: IPR003960; AAA_sub.
DR Pfam: PF00004; AAA; 2.
DR PROSITE: PS00674; AAA; UNKNOWN_1.
SQ SEQUENCE 785 AA; 88003 MW; DB75ED37E13AFC63 CRC64;

Query Match
Best Local Similarity 59.1%; Score 39; DB 5; Length 785;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFK 11
   |:| | | | | | |
Db 591 AKSASFHTKF 601

RESULT 12
Q91502 PRELIMINARY; PRT; 924 AA.
AC Q91502;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Organic solvent tolerance protein OstA precursor.
GN OSTA OR PA0595.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAOI;
RX MEDLINE=20437337; PubMed=10984043;
RA Hickey C.K., Pham X.-Q.T., Erwin A.L., Mizouchi S.D., Warren P.,
RA Stover M.J., Brinkman F.S.D., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong K.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAOI, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL: AE004495; AAG03984.1; -.
KW Complete proteome.
SQ SEQUENCE 924 AA; 104271 MW; 87F58C83E218335B CRC64;

Query Match
Best Local Similarity 59.1%; Score 39; DB 16; Length 924;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 SLSEFXFTKFDLD 14
   |:|:| | | | |
Db 569 TLKYLYTRYDLD 580

RESULT 13
Q43376 PRELIMINARY; PRT; 254 AA.
AC Q43376;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Mannose/glucose-binding lectin precursor (Fragment).
GN LEC.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV- SELLIE; TISSUE=SEED;
RA Law I.J.;

*Cloning and expression of cDNA for mannose-binding lectin from
peanut.*;
RT Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RL EMBL: U22472; AAA74575.1; -.
DR HSP: P02867; 2BQP.
DR InterPro: IPR000985; Lectin_legA.
DR InterPro: IPR001220; Lectin_legB.
DR Pfam: PF00138; lectin_legA; 1.
DR Pfam: PF00139; lectin_legB; 1.
DR ProDom: PD000671; Lectin_legA; 1.
DR ProDom: PD000711; Lectin_legB; 1.
DR PROSITE: PS00307; LECTIN_LEGWE_BETA; UNKNOWN_1.
KW Lectin; Signal.
FT NON_TER 1
FT SIGNAL <1 2
SQ SEQUENCE 254 AA; 28202 MW; 3AE1EA1F90B1CA03 CRC64;

Query Match
Best Local Similarity 57.6%; Score 38; DB 10; Length 254;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSEFXFTKFDLD 14
   |:|:| | | | |
Db 3 SLSESYNKFQD 14

RESULT 14
Q20265 PRELIMINARY; PRT; 467 AA.
AC Q20265;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE F41C3.2 protein.
GN F41C3.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Willson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX Chisoe S.;
RT "The sequence of C. elegans cosmid F41C3.";
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL: U23521; AAC46809.1; -.
DR InterPro: IPR003662; subtransporter.
DR Pfam: PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 467 AA; 51868 MW; D15F5181F0096C67 CRC64;

```

Query Match 57.6%; Score 38; DB 5; Length 467;  
 Best Local Similarity 58.3%; Pred. No. 34;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDLD 14  
 ||| |::| |  
 Db 52 SLNFNFSKFEQD 63

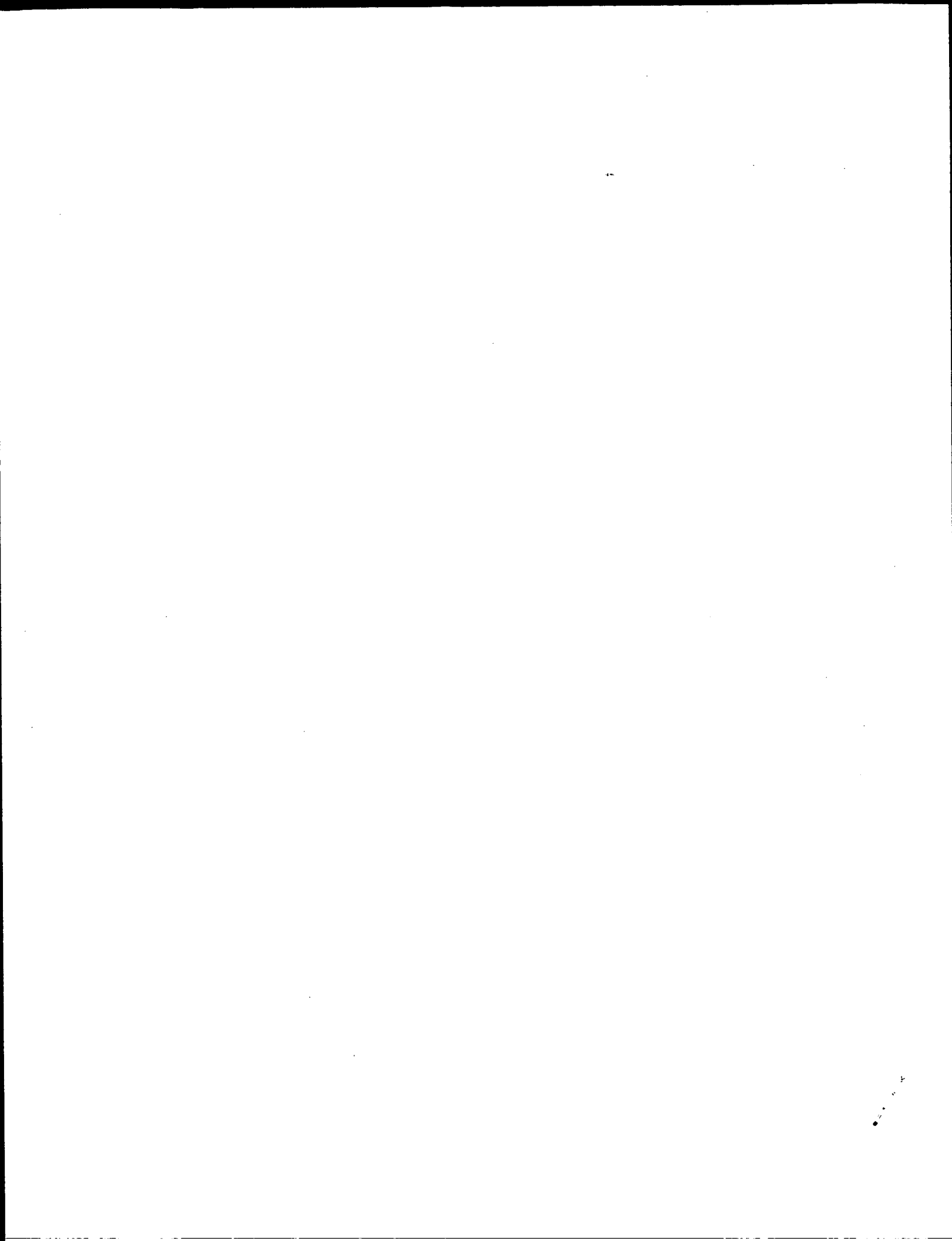
## RESULT 15

Q23024 PRELIMINARY; PRT; 705 AA.  
 AC Q23024; Q23025;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
 DE UNC-93 protein.  
 GN UNC-93.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=92210636; PubMed=1313436;  
 RA Levin J.Z., Horvitz R.H.;  
 RT "The Caenorhabditis elegans unc-93 gene encodes a putative  
 transmembrane protein that regulates muscle contraction.";  
 RL J. Cell Biol. 117:143-155(1992).  
 DR EMBL; X64415; CAA45760.1; -;  
 DR EMBL; X64415; CAA45761.1; -;  
 SQ SEQUENCE 705 AA; 80351 MW; 91D49A788A3EAC58 CRC64;

Query Match 57.6%; Score 38; DB 5; Length 705;  
 Best Local Similarity 70.0%; Pred. No. 52;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 QSLSFXTKF 11  
 ||| |::| |  
 Db 623 QSLQFAFTKY 632

Search completed: February 26, 2003, 14:51:50  
 Job time : 31 secs





XX PS Claim 4; Page 28; 46pp; English.

CC CC The peptides AAW61497-W61502 can be used to form pylartin proteins which

CC CC preserve progenitor cells (pc). The proteins can be used for protecting

CC CC the integrity of the hematopoietic processes in vivo and as adjuncts in

CC CC therapeutic treatments related to cancer and other diseases which can

CC CC otherwise adversely impact upon the hematopoietic system. Since the

CC CC proteins bind specifically to primitive PCs, they can also be used for

CC CC the identification and localisation of PCs. The methods can be used for

CC CC e.g. expanding PC populations ex vivo to increase chances of

CC CC engraftation. Improving conditions for transporting and storing PCs and

CC CC for removing a fundamental barrier thereby enabling gene therapy to

CC CC treat and cure a broad range of life-threatening haematologic diseases

CC CC such as sickle cell anaemia and thalassemia.

XX SQ Sequence 14 AA;

Query Match 98.5%; Score 65; DB 19; Length 14;

Best Local Similarity 92.9%; Pred. No. 9.1e-06;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14

DB 1 AQSLSFSFTKFDLD 14

RESULT 2

AA67818

ID AAR67818 standard; peptide; 14 AA.

XX AC AAR67818;

XX DT 18-AUG-1995 (first entry)

XX DE Flk2 ligand N-terminal sequence.

XX KW human Flk2 receptor protein-tyrosine-kinase ligand; peripheral

XX KW peripheral blood leukocyte conditioned medium;

XX KW bone marrow disorder diagnosis; hematopoietic stem cell;

XX KW proliferation; differentiation.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT misc\_difference 7 /note= "any amino acid"

FT FT

XX PN W09500554-A.

XX PD 05-JAN-1995.

XX PF 17-JUN-1994; 94WO-US06944.

XX PR 18-JUN-1993; 93US-0080244.

XX PR 21-JUN-1993; 93US-0081508.

XX PR 23-NOV-1993; 93US-0157490.

XX PA (UYPR-) UNIV PRINCETON.

XX PI Lemischka IR;

XX DR WPI; 1995-052014/07.

XX PT Ligand for receptor protein tyrosine kinase - useful for the

XX PT stimulation of primitive hematopoietic stem cells causing

XX PT proliferation and/or differentiation

XX PS Claim 1; Page 105; 131pp; English.

XX CC The sequence corresponds to the N-terminal region of a human Flk2

CC CC receptor protein-tyrosine-kinase ligand, isolated from

CC CC phytohemagglutinin-stimulated human peripheral blood leukocyte

CC tissue culture conditioned medium. The ligand may be used in

CC diagnosis of bone marrow disorders, and to stimulate the

CC proliferation and/or differentiation of primitive hematopoietic stem

CC cells. The ligand binds to a receptor protein-tyrosine-kinase

CC expressed in primitive but not mature mammalian hematopoietic cells.

XX SQ Sequence 14 AA;

Query Match 97.0%; Score 64; DB 16; Length 14;

Best Local Similarity 100.0%; Pred. No. 1.4e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14

DB 1 AQSLSFXFTKFDLD 14

RESULT 3

AA62895

ID AAG62895 standard; peptide; 14 AA.

XX AC AAG62895;

XX DT 17-SEP-2001 (first entry)

XX DE Antigenic peptide derived from a french bean FRIL polypeptide.

XX KW FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;

XX KW progenitor cell preservation factor; radiotherapy; chemotherapy;

XX KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;

XX KW severe combined immunodeficiency; aplastic anemia; tissue repair.

XX OS Phaseolus vulgaris.

XX FH Key Location/Qualifiers

FT Misc-difference 7 /label= Asn, Cys, Ser

FT FT

XX PN W0200149851-A1.

XX PD 12-JUL-2001.

XX PF 30-DEC-1999; 99WO-US31307.

XX PR 30-DEC-1999; 99WO-US31307.

XX PA (PHYL-) PHYLOGIX LLC.

XX PI Colucci MG, Chrispeels MJ, Moore JG;

XX DR WPI; 2001-441882/47.

XX PT Legume progenitor cell preservation factors for in vivo or ex vivo

XX PT preservation of hematopoietic progenitor cells and as therapeutics for

XX PT alleviating/reducing progenitor cell-depleting activity of cancer

XX PT therapeutics

XX PS Example 5; Page 72; 173pp; English.

XX CC The present sequence is derived from FRIL (Flk2/Flt3 tyrosine kinase

CC CC receptor-interacting lectin), and is used to raise antibodies. The

CC CC specification describes a composition of one or more members

CC CC of FRIL family of progenitor cell preservation factors. The composition

CC CC is useful for alleviating or reducing the hematopoietic progenitor

CC CC cell-depleting activity of a therapeutic treatment, including

CC CC radiotherapeutic and/or chemotherapeutic treatments. Administration of

CC CC FRIL compositions to a patient prior to treatment of the patient with

CC CC a therapeutic treatment having a hematopoietic progenitor cell-depleting

CC CC activity alleviates or reduces the hematopoietic progenitor

CC CC cell-depleting activity of the therapeutic treatment in the patient.

CC CC FRIL family members are useful for isolating population of progenitor

CC CC cells, hemangioblasts, and mesenchymal stem cells. The composition is

CC CC administered to reduce progenitor cell depleting effects of



CC chemotherapeutics, so that the patient can receive a higher dose of the  
 CC chemotherapeutic and preferably recover from cancer. It is also  
 CC administered to patients having, or predisposed to developing a  
 CC condition where the patients hematopoietic progenitor cells are  
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
 CC The isolated mesenchymal cells are useful for tissue repair.

XX  
 SQ Sequence 14 AA;

Query Match 97.0%; Score 64; DB 22; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14  
 |||||  
 Db 1 AQSLSFXFTKFDLD 14

# RESULT 4

AAG62898  
 ID AAG62898 standard; Protein; 303 AA.

AC AAG62898;

DT 17-SEP-2001 (first entry)

XX Amino acid sequence of a french bean FRIL polypeptide.

XX FRIL; FLK2/Flt3 tyrosine kinase receptor-interacting lectin;  
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;  
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
 KW severe combined immunodeficiency; aplastic anemia; tissue repair.

XX Phaseolus vulgaris.

XX WO200149851-A1.

XX 12-JUL-2001.

XX 30-DEC-1999; 99WO-US31307.

XX 30-DEC-1999; 99WO-US31307.

XX (PHYL-) PHYLOGIX LLC.

XX Colucci MG, Chrispeels MJ, Moore JG;

XX WPI; 2001-441882/47.

XX N-PSDB; AAH42306.

XX Legume progenitor cell preservation factors for in vivo or ex vivo  
 PT preservation of hematopoietic progenitor cells and as therapeutics for  
 PT alleviating/reducing progenitor cell-depleting activity of cancer  
 PT therapeutics -

XX Example 5; Page 81; 173pp; English.

XX The present sequence represents a FRIL (FLK2/Flt3 tyrosine kinase  
 CC receptor-interacting lectin) polypeptide. The specification describes a  
 CC composition of one or more members of FRIL family of progenitor cell  
 CC preservation factors. The composition is useful for alleviating or  
 CC reducing the hematopoietic progenitor cell-depleting activity of  
 CC a therapeutic treatment, including radiotherapeutic and/or  
 CC chemotherapeutic treatments. Administration of FRIL compositions to a  
 CC patient prior to treatment of the patient with a therapeutic treatment  
 CC having a hematopoietic progenitor cell-depleting activity alleviates or  
 CC reduces the hematopoietic progenitor cell-depleting activity of the  
 CC therapeutic treatment in the patient. FRIL family members are useful for  
 CC isolating population of progenitor cells, hemangioblasts, and mesenchymal  
 CC stem cells. The composition is administered to reduce progenitor cell  
 CC depleting effects of chemotherapeutics, so that the patient can receive  
 CC a higher dose of the chemotherapeutic and preferably recover from cancer.  
 CC It is also administered to patients having, or predisposed to developing

CC a condition where the patients hematopoietic progenitor cells are  
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
 CC The isolated mesenchymal cells are useful for tissue repair.

XX Sequence 303 AA;

Query Match 97.0%; Score 64; DB 22; Length 303;  
 Best Local Similarity 92.9%; Pred. No. 0.0004;  
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14  
 |||||  
 Db 1 AQSLSFNFTKFDLD 14

# RESULT 5

AAG62899

ID AAG62899 standard; peptide; 15 AA.

XX AAG62899;

XX 17-SEP-2001 (first entry)

XX Peptide derived from a yam FRIL polypeptide.

XX FRIL; FLK2/Flt3 tyrosine kinase receptor-interacting lectin;  
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;  
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
 KW severe combined immunodeficiency; aplastic anemia; tissue repair.

XX Sphenostylis stenocarpa.

XX WO200149851-A1.

XX 12-JUL-2001.

XX 30-DEC-1999; 99WO-US31307.

XX 30-DEC-1999; 99WO-US31307.

XX (PHYL-) PHYLOGIX LLC.

XX Colucci MG, Chrispeels MJ, Moore JG;

XX WPI; 2001-441882/47.

XX Legume progenitor cell preservation factors for in vivo or ex vivo  
 PT preservation of hematopoietic progenitor cells and as therapeutics for  
 PT alleviating/reducing progenitor cell-depleting activity of cancer  
 PT therapeutics -

XX Example 22; Page 116; 173pp; English.

XX The present sequence is derived from a FRIL (FLK2/Flt3 tyrosine kinase  
 CC receptor-interacting lectin) polypeptide. The specification describes a  
 CC composition of one or more members of FRIL family of progenitor cell  
 CC preservation factors. The composition is useful for alleviating or  
 CC reducing the hematopoietic progenitor cell-depleting activity of a  
 CC therapeutic treatment, including radiotherapeutic and/or  
 CC chemotherapeutic treatments. Administration of FRIL compositions to a  
 CC patient prior to treatment of the patient with a therapeutic treatment  
 CC having a hematopoietic progenitor cell-depleting activity alleviates or  
 CC reduces the hematopoietic progenitor cell-depleting activity of the  
 CC therapeutic treatment in the patient. FRIL family members are useful for  
 CC isolating population of progenitor cells, hemangioblasts, and mesenchymal  
 CC stem cells. The composition is administered to reduce progenitor cell  
 CC depleting effects of chemotherapeutics, so that the patient can receive  
 CC a higher dose of the chemotherapeutic and preferably recover from cancer.  
 CC It is also administered to patients having, or predisposed to developing  
 CC a condition where the patients hematopoietic progenitor cells are  
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
 CC The isolated mesenchymal cells are useful for tissue repair.

SQ Sequence 15 AA; Query Match 84.8%; Score 56; DB 22; Length 15;  
 Best Local Similarity 78.6%; Pred. No. 0.0005; Mismatches 2; Indels 0; Gaps 0;  
 Matches 11; Conservative 1;

QY 1 AQSLSFXFTKFDLD 14  
 |||:| | | | | |  
 Db 1 AQSVSFTTKFDSD 14

RESULT 6  
 AAG62901  
 ID AAG62901 standard; Protein; 234 AA.  
 XX AC AAG62901;  
 XX DT 17-SEP-2001 (first entry)  
 XX DE Amino acid sequence of a partial FRIL polypeptide.  
 XX KW FRIL; FLK2/Flt3 tyrosine kinase receptor-interacting lectin;  
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;  
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
 KW severe combined immunodeficiency; aplastic anemia; tissue repair.  
 XX OS Sphenostylis stenocarpa.  
 XX PN WO200149851-A1.  
 XX PD 12-JUL-2001.  
 XX PF 30-DEC-1999; 99WO-US31307.  
 XX PR 30-DEC-1999; 99WO-US31307.  
 XX PA (PHYL-) PHYLOGIX LLC.  
 XX PI Colucci MG, Chrispeels MJ, Moore JG;  
 XX DR WPI; 2001-441882/47.  
 XX Legume Progenitor cell preservation factors for in vivo or ex vivo  
 PT preservation of hematopoietic progenitor cells and as therapeutics for  
 PT alleviating/reducing progenitor cell-depleting activity of cancer  
 PT therapeutics -  
 XX Example 22; Page 116-117; 173pp; English.

The present sequence is a partial a FRIL (Flk2/Flt3 tyrosine kinase  
 receptor-interacting lectin) protein. The specification describes a  
 composition of one or more members of FRIL family of progenitor cell  
 preservation factors. The composition is useful for alleviating or  
 reducing the hematopoietic progenitor cell-depleting activity of a  
 therapeutic treatment, including radiotherapeutic and/or  
 chemotherapeutic treatments. Administration of FRIL compositions to a  
 patient prior to treatment of the patient with a therapeutic treatment  
 reduces the hematopoietic progenitor cell-depleting activity of the  
 therapeutic treatment in the patient. FRIL family members are useful for  
 isolating population of progenitor cells, hemangioblasts, and mesenchymal  
 stem cells. The composition is administered to reduce progenitor cell  
 depleting effects of the chemotherapeutic, so that the patient can receive  
 a higher dose of the chemotherapeutic and preferably recover from cancer.  
 It is also administered to patients having, or predisposed to developing  
 a condition where the patients hematopoietic progenitor cells are  
 depleted, such as severe combined immunodeficiency or aplastic anemia.  
 The isolated mesenchymal cells are useful for tissue repair.

SQ Sequence 234 AA;  
 Query Match 84.8%; Score 56; DB 22; Length 234;  
 Best Local Similarity 78.6%; Pred. No. 0.0099;

Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14  
 |||:| | | | | |  
 Db 1 AQSVSFTTKFDSD 14

RESULT 7  
 AAW61497  
 ID AAW61497 standard; peptide; 12 AA.  
 XX AC AAW61497;  
 XX DT 19-OCT-1998 (first entry)  
 XX DE Pylartin protein, peptide chain beta (i).  
 XX KW Pylartin protein; progenitor cell; hematopoietic system; cancer;  
 KW engraftation; hematologic disease; sickle cell anemia; thalassemia.  
 XX OS Leguminosae.  
 XX PN WO9825457-A1.  
 XX PD 18-JUN-1998.  
 XX PF 09-DEC-1997; 97WO-US22486.  
 XX PR 28-MAR-1997; 97US-0825369.  
 XX PR 09-DEC-1996; 96US-0762537.  
 XX PA (INCL-) IMCLONE SYSTEMS INC.  
 XX PI Moore JG;  
 XX DR WPI; 1998-348161/30.  
 XX Pylartin proteins which preserve progenitor cells - useful for  
 PT hematopoietic therapies in cancer treatment or for treating  
 PT hematologic diseases  
 XX Claim 1; Page 28; 46pp; English.

The peptides AAW61497-W61502 can be used to form pylartin proteins which  
 preserve progenitor cells (pc). The proteins can be used for protecting  
 the integrity of the hematopoietic processes in vivo and as adjuncts in  
 therapeutic treatments related to cancer and other diseases which can  
 otherwise adversely impact upon the hematopoietic system. Since the  
 proteins bind specifically to primitive PCs, they can also be used for  
 the identification and localisation of PCs. The methods can be used for  
 e.g. expanding PC populations ex vivo to increase chances of  
 engraftation, improving conditions for transporting and storing PCs and  
 for removing a fundamental barrier thereby enabling gene therapy to  
 treat and cure a broad range of life-threatening hematologic diseases  
 such as sickle cell anaemia and thalassemia.

SQ Sequence 12 AA;  
 Query Match 83.3%; Score 55; DB 19; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 0.00061;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12  
 ||| | | | | |  
 Db 1 AQSLSFXFTKFD 12

RESULT 8  
 AAW87973  
 ID AAW87973 standard; Protein; 264 AA.  
 XX AC AAW87973;  
 XX

DT 13-APR-1999 (first entry)  
 DE A lectin derived progenitor cell preservation factor.  
 XX  
 XX Lactin derived progenitor cell preservation factor; progenitor cell;  
 KW haematopoietic cell; cultured cell preservation; anticancer therapy;  
 KW myeloablative therapy; sickle-cell anaemia; ablative therapy protection;  
 KW FLK2/FLT3 receptor.  
 XX  
 XX Dolichos lab lab.  
 OS  
 XX  
 XX WO9859038-A1.  
 PN  
 XX  
 XX 30-DEC-1998.  
 PD  
 XX  
 XX 23-JUN-1998; 98WO-US13046.  
 PF  
 XX  
 XX 24-JUN-1997; 97US-0881189.  
 PR  
 XX  
 XX (IMCL-) IMCLONE SYSTEMS INC.  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 XX Chrispeels MJ, Colucci MG, Moore JG;  
 PI  
 XX  
 XX WPI; 1999-081274/07.  
 DR  
 XX  
 XX N-PSDB; AAX03593.  
 XX  
 XX New nucleic acid encoding plant lectin that preserves progenitor  
 PT cells - particularly haematopoietic progenitors, useful for bone  
 PT marrow reconstitution after ablative therapy, and to increase DNA  
 PT transfer in gene therapy  
 XX  
 XX Claim 1; Page 30-31; 72pp; English.  
 PS  
 XX The present sequence represents a lectin derived progenitor cell  
 CC preservation factor. The protein is used to preserve unipotent,  
 CC pluripotent or totipotent progenitor cells, especially haematopoietic  
 CC cells, and also progenitors from nerve, muscle, skin, gut, bone,  
 CC kidney, liver, pancreas or thymus. Specific applications are  
 CC preservation of cultured cells intended for administration after  
 CC (anticancer) myeloablative therapy (bone marrow or whole-body irradiation  
 CC or chemotherapy) to reconstitute the haematopoietic system; enrichment  
 CC of progenitor cells (e.g. during ex vivo purging of malignant cells);  
 CC treatment of tissues containing haematopoietic progenitors for subsequent  
 CC transplant to improve haematopoietic competence; improving transfer of  
 CC exogenous DNA to progenitor cells (in gene therapy of various  
 CC haematological disorders, e.g. sickle-cell anaemia); and protection  
 CC against ablative therapy (to eliminate proliferating cells specifically),  
 CC followed by re-establishment of differentiation and proliferation of  
 CC preserved progenitors. The protein, when linked to magnetic beads, may  
 CC also be used to isolate cells that express the FLK2/FLT3 receptor.  
 XX  
 XX SQ Sequence 264 AA;

Query Match 83.3%; Score 55; DB 20; Length 264;  
 Best Local Similarity 91.7%; Pred. No. 0.017;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12  
 ||||| |||||  
 Db 1 AQSLSFSFTKFD 12

RESULT 9  
 AAG62890  
 ID AAG62890 standard; Protein; 264 AA.  
 XX  
 XX AAG62890;  
 AC

DT 17-SEP-2001 (first entry)

DE Amino acid sequence of a hyacinth bean FRIL polypeptide.

XX

KW FRIL; FLK2/Flt3 tyrosine kinase receptor-interacting lectin;  
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;  
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
 KW severe combined immunodeficiency; aplastic anemia; tissue repair.  
 XX  
 XX Dolichos lab lab.  
 OS  
 XX WO200149851-A1.  
 PN  
 XX  
 XX 12-JUL-2001.  
 PD  
 XX  
 XX 30-DEC-1999; 99WO-US31307.  
 PF  
 XX  
 XX 30-DEC-1999; 99WO-US31307.  
 PR  
 XX  
 XX (PHYL-) PHYLOGIX LLC.  
 PA  
 XX  
 XX Colucci MG, Chrispeels MJ, Moore JG;  
 PI  
 XX  
 XX WPI; 2001-441882/47.  
 DR  
 XX  
 XX N-PSDB; AAH42287.  
 XX  
 XX Legume progenitor cell preservation factors for in vivo or ex vivo  
 PT preservation of hematopoietic progenitor cells and as therapeutics for  
 PT alleviating/reducing progenitor cell-depleting activity of cancer  
 PT therapeutics  
 XX  
 XX Example 1; Page 54-55; 173pp; English.  
 PS  
 XX The present sequence represents a FRIL (FLK2/Flt3 tyrosine kinase  
 CC receptor-interacting lectin) polypeptide. The specification describes a  
 CC composition of one or more members of FRIL family of progenitor cell  
 CC preservation factors. The composition is useful for alleviating or  
 CC reducing the hematopoietic progenitor cell-depleting activity of  
 CC a therapeutic treatment, including radiotherapeutic and/or  
 CC chemotherapeutic treatments. Administration of FRIL compositions to a  
 CC patient prior to treatment of the patient with a therapeutic treatment  
 CC having a hematopoietic progenitor cell-depleting activity alleviates or  
 CC reduces the hematopoietic progenitor cell-depleting activity of the  
 CC therapeutic treatment in the patient. FRIL family members are useful for  
 CC isolating population of progenitor cells, hemangioblasts, and mesenchymal  
 CC stem cells. The composition is administered to reduce progenitor cell  
 CC depleting effects of the chemotherapeutic, so that the patient can receive  
 CC a higher dose of the chemotherapeutic and preferably recover from cancer.  
 CC It is also administered to patients having, or predisposed to developing  
 CC a condition where the patients hematopoietic progenitor cells are  
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
 CC The isolated mesenchymal cells are useful for tissue repair.

SQ Sequence 264 AA;

Query Match 83.3%; Score 55; DB 22; Length 264;  
 Best Local Similarity 91.7%; Pred. No. 0.017;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12  
 ||||| |||||  
 Db 1 AQSLSFSFTKFD 12

RESULT 10  
 AAG62894  
 ID AAG62894 standard; Protein; 286 AA.  
 XX

AC AAG62894;

XX 17-SEP-2001 (first entry)

DE Alpha-amylase inhibitor signal peptide and FRIL fusion.

KW FRIL; FLK2/Flt3 tyrosine kinase receptor-interacting lectin;  
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;  
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;

KW severe combined immunodeficiency; aplastic anemia; tissue repair;  
 KW alpha-amylase inhibitor gene.

XX Synthetic.  
 OS Unidentified.  
 OS Dolichos lab lab.  
 XX WO200149851-A1.  
 XX 12-JUL-2001.

XX 30-DEC-1999; 99WO-US31307.  
 XX 30-DEC-1999; 99WO-US31307.  
 XX (PHYL-) PHYLOGIX LLC.

XX Colucci MG, Chrispeels MJ, Moore JG;  
 XX WPI; 2001-441882/47.  
 XX N-PSDB; AAH42295.

XX Legume Progenitor cell preservation factors for in vivo or ex vivo  
 PT preservation of hematopoietic progenitor cells and as therapeutics for  
 PT alleviating/reducing progenitor cell-depleting activity of cancer  
 PT therapeutics -

XX Example 1; Page 59; 173pp; English.

XX The present sequence represents fusion protein of alpha-amylase inhibitor  
 CC signal peptide and FRIL (Flk2/Flt3 tyrosine kinase receptor-interacting  
 CC lectin). The specification describes a composition of one or more members  
 CC of FRIL family of progenitor cell preservation factors. The composition  
 CC is useful for alleviating or reducing the hematopoietic progenitor  
 CC cell-depleting activity of a therapeutic treatment, including  
 CC radiotherapeutic and/or chemotherapeutic treatments. Administration of  
 CC FRIL compositions to a patient prior to treatment of the patient with  
 CC a therapeutic treatment having a hematopoietic progenitor cell-depleting  
 CC activity alleviates or reduces the hematopoietic progenitor  
 CC cell-depleting activity of the therapeutic treatment in the patient.  
 CC FRIL family members are useful for isolating population of progenitor  
 CC cells, hemangioblasts, and mesenchymal stem cells. The composition is  
 CC administered to reduce progenitor cell depleting effects of  
 CC chemotherapeutics, so that the patient can receive a higher dose of the  
 CC chemotherapeutic and preferably recover from cancer. It is also  
 CC administered to patients having, or predisposed to developing a  
 CC condition where the patients hematopoietic progenitor cells are  
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
 CC The isolated mesenchymal cells are useful for tissue repair.

XX Sequence 286 AA;

Query Match 83.3%; Score 55; DB 22; Length 286;  
 Best Local Similarity 91.7%; Pred. No. 0.019;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXTKFD 12  
 ||||| |||||  
 Db 23 AQSLSFSFTKFD 34

RESULT 11  
 AAG62896  
 ID AAG62896 standard; peptide; 14 AA.  
 XX AC AAG62896;

XX 17-SEP-2001 (first entry)  
 XX Peptide derived from a french bean FRIL polypeptide.  
 DE FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;  
 XX progenitor cell preservation factor; radiotherapy; chemotherapy;  
 KW

KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
 KW severe combined immunodeficiency; aplastic anemia; tissue repair.  
 XX Phaseolus vulgaris.

XX Key Location/Qualifiers  
 FT Misc-difference 7 /note= "not known"  
 FT

XX WO200149851-A1.

XX 12-JUL-2001.

XX 30-DEC-1999; 99WO-US31307.

XX 30-DEC-1999; 99WO-US31307.

XX (PHYL-) PHYLOGIX LLC.

XX Colucci MG, Chrispeels MJ, Moore JG;

XX WPI; 2001-441882/47.

XX Legume Progenitor cell preservation factors for in vivo or ex vivo  
 PT preservation of hematopoietic progenitor cells and as therapeutics for  
 PT alleviating/reducing progenitor cell-depleting activity of cancer  
 PT therapeutics -

XX Example 5; Page 75; 173pp; English.

XX The present sequence is derived from a FRIL (Flk2/Flt3 tyrosine kinase  
 CC receptor-interacting lectin) polypeptide. The specification describes a  
 CC composition of one or more members of FRIL family of progenitor cell  
 CC preservation factors. The composition is useful for alleviating or  
 CC reducing the hematopoietic progenitor cell-depleting activity of a  
 CC therapeutic treatment, including radiotherapeutic and/or  
 CC chemotherapeutic treatments. Administration of FRIL compositions to a  
 CC patient prior to treatment of the patient with a therapeutic treatment  
 CC having a hematopoietic progenitor cell-depleting activity alleviates or  
 CC reduces the hematopoietic progenitor cell-depleting activity of the  
 CC therapeutic treatment in the patient. FRIL family members are useful for  
 CC isolating population of progenitor cells, hemangioblasts, and mesenchymal  
 CC stem cells. The composition is administered to reduce progenitor cell  
 CC depleting effects of chemotherapeutics, so that the patient can receive  
 CC a higher dose of the chemotherapeutic and preferably recover from cancer.  
 CC It is also administered to patients having, or predisposed to developing  
 CC a condition where the patients hematopoietic progenitor cells are  
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
 CC The isolated mesenchymal cells are useful for tissue repair.

XX Sequence 14 AA;

Query Match 71.2%; Score 47; DB 22; Length 14;  
 Best Local Similarity 85.7%; Pred. No. 0.024;  
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSFXTKFDLD 14  
 ||||| |||||  
 Db 1 AQSLSFXTKDALD 14

RESULT 12  
 ABB68966  
 ID ABB68966 standard; Protein; 896 AA.  
 XX AC ABB68966;

XX 26-MAR-2002 (first entry)  
 DT Drosophila melanogaster polypeptide SEQ ID NO 33690.  
 DE Drosophila; developmental biology; cell signalling; insecticide;  
 XX pharmaceutical.  
 KW

XX OS Drosophila melanogaster.  
XX PN WO200171042-A2.  
XX PD 27-SEP-2001.  
XX PF 23-MAR-2001; 2001WO-US09231.  
XX PR 23-MAR-2000; 2000US-191637P.  
XX PR 11-JUL-2000; 2000US-0614150.  
XX PA (PEKE ) PE CORP NY.  
XX PI Venter JC, Adams M, Li PWD, Myers EW;  
XX DR WPI: 2001-656860/75.  
XX DR N-PSDB; ABL13069.  
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell  
XX PT interactions -  
XX PS Disclosure; SEQ ID NO 33690; 21pp + Sequence Listing; English.  
XX CC The invention relates to an isolated nucleic acid detection reagent  
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is  
XX CC useful in developmental biology and in elucidating cell signalling and  
XX CC cell-cell interactions in higher eukaryotes for the development of  
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention  
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
XX CC sequences (ABL01840-ABL16175) and the encoded proteins  
XX CC (ABB57737-ABB72072).  
XX CC The sequence data for this patent did not form part of the printed  
XX CC specification, but was obtained in electronic format directly from WIPO  
XX CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 896 AA;  
  
Query Match 63.6%; Score 42; DB 22; Length 896;  
Best Local Similarity 57.1%; Pred. No. 19;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 AQSLSFXFTKFDLD 14  
: | | | | :  
Db 158 SQPLQFSFTFYDLD 171  
  
RESULT 13  
AAB08214  
ID AAB08214 standard; Protein; 928 AA.  
XX AC AAB08214;  
XX DT 04-DEC-2000 (first entry)  
XX DE Amino acid sequence of a Drosophila naked cuticle polypeptide.  
XX KW Nkd gene; naked cuticle gene; segment-polarity gene; Wnt signalling;  
XX KW gene therapy; Nkd defect; cancer.  
XX OS Drosophila melanogaster.  
XX PN WO200049034-A1.  
XX PD 24-AUG-2000.  
XX PF 17-FEB-2000; 2000WO-US04188.  
XX PR 17-FEB-1999; 99US-0120646.  
XX PA (STRD ) UNIV LELAND STANFORD JUNIOR.  
XX PR

PI Scott M, Zeng W, Wharton K;  
XX WPI: 2000-571967/53.  
DR N-PSDB; AAA63923.  
XX An isolated nucleic acid molecule useful for analyzing (genetic  
XX PT predisposition to) a disease state and for therapeutic purposes e.g.  
XX PT treatment of cancer comprises a sequence encoding a naked cuticle  
XX PT protein -  
XX PS Claim 12; Page 41-43; 58pp; English.  
XX CC The present sequence represents a Nkd (naked cuticle) polypeptide. In  
XX CC Drosophila, Nkd is a segment-polarity gene whose expression is  
XX CC induced by Wnt signalling. The Nkd polypeptide acts to antagonize Wnt  
XX CC signalling. Nkd may link ion fluxes to the regulation of Wnt signal  
XX CC potency, duration or distribution. The Nkd polynucleotides can be  
XX CC used for identifying homologous or related proteins, to modulate the  
XX CC expression or function of Nkd polypeptides, and in studying associated  
XX CC physiological pathways. Nkd polynucleotides can also be used in gene  
XX CC therapy to treat disorders associated with Nkd defects. They may  
XX CC also be used for therapeutic purposes e.g. treatment of cancer.  
XX SQ Sequence 928 AA;  
  
Query Match 63.6%; Score 42; DB 21; Length 928;  
Best Local Similarity 57.1%; Pred. No. 20;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 AQSLSFXFTKFDLD 14  
: | | | | :  
Db 190 SQPLQFSFTFYDLD 203  
  
RESULT 14  
AAG56205  
ID AAG56205 standard; Protein; 44 AA.  
XX AC AAG56205;  
XX DT 18-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 72209.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX PR 25-FEB-1999; 99US-0121825.  
XX PR 05-MAR-1999; 99US-0123180.  
XX PR 09-MAR-1999; 99US-0123548.  
XX PR 23-MAR-1999; 99US-0125788.  
XX PR 29-MAR-1999; 99US-0126264.  
XX PR 29-MAR-1999; 99US-0126785.  
XX PR 01-APR-1999; 99US-0127462.  
XX PR 06-APR-1999; 99US-0128234.  
XX PR 08-APR-1999; 99US-0128714.  
XX PR 16-APR-1999; 99US-0129845.  
XX PR 19-APR-1999; 99US-0130077.  
XX PR 21-APR-1999; 99US-0130449.  
XX PR 23-APR-1999; 99US-0130510.  
XX PR 28-APR-1999; 99US-0130891.  
XX PR 30-APR-1999; 99US-0131449.  
XX PR 30-APR-1999; 99US-0132048.  
XX PR 30-APR-1999; 99US-0132407.



PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161922.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 62.1%; Score 41; DB 21; Length 44;  
 Best Local Similarity 61.5%; Pred. NO. 1.1;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 OSLSEFTKEDLD 14  
 :||||| |||  
 Db 10 ESLSLFVGDFED 22

## RESULT 15

AAG60788  
 ID AAG60788 standard; Protein: 44 AA.

XX AC AAG60788;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 78778.

XX KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 03-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127482.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132853.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 21-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139452.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 31-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0156559.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 23-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 62.1%; Score 41; DB 21; Length 44;  
Best Local Similarity 61.5%; Pred. No. 1.1;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 QSLSEFTKFDLD 14  
Db 10 ELSLFLVGFDFD 22

Search completed: February 26, 2003, 14:50:53  
Job time : 84 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:51:00 ; Search time 12 Seconds  
(without alignments)  
44.009 Million cell updates/sec

Title: US-09-476-485A-31  
Perfect score: 66  
Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 174566 seqs, 37721826 residues

Total number of hits satisfying chosen parameters: 174566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 65    | 98.5        | 14     | 10 | US-09-934-251A-2  |
| 2          | 55    | 83.3        | 12     | 10 | US-09-934-251A-1  |
| 3          | 42    | 63.6        | 60     | 10 | US-09-730-989-22  |
| 4          | 42    | 63.6        | 668    | 10 | US-09-730-989-20  |
| 5          | 37    | 56.1        | 73     | 9  | US-10-063-547-148 |
| 6          | 37    | 56.1        | 73     | 9  | US-10-174-590-396 |
| 7          | 37    | 56.1        | 73     | 9  | US-10-176-758-396 |
| 8          | 37    | 56.1        | 73     | 9  | US-10-063-616-148 |
| 9          | 37    | 56.1        | 73     | 9  | US-10-175-737-396 |
| 10         | 37    | 56.1        | 73     | 9  | US-10-063-502-148 |
| 11         | 37    | 56.1        | 73     | 9  | US-10-173-706-396 |
| 12         | 37    | 56.1        | 73     | 9  | US-10-175-738-396 |
| 13         | 37    | 56.1        | 73     | 9  | US-10-175-752-396 |
| 14         | 37    | 56.1        | 73     | 9  | US-10-176-482-396 |
| 15         | 37    | 56.1        | 73     | 9  | US-10-176-757-396 |
| 16         | 37    | 56.1        | 73     | 9  | US-10-176-913-396 |
| 17         | 37    | 56.1        | 73     | 9  | US-10-180-552-396 |
| 18         | 37    | 56.1        | 73     | 9  | US-10-180-557-396 |
| 19         | 37    | 56.1        | 73     | 9  | US-10-173-700-396 |

20 37 56.1 73 9 US-10-174-572-396 Sequence 396, App  
21 37 56.1 73 9 US-10-174-579-396 Sequence 396, App  
22 37 56.1 73 9 US-10-174-582-396 Sequence 396, App  
23 37 56.1 73 9 US-10-174-588-396 Sequence 396, App  
24 37 56.1 73 9 US-10-175-739-396 Sequence 396, App  
25 37 56.1 73 9 US-10-175-740-396 Sequence 396, App  
26 37 56.1 73 9 US-10-175-743-396 Sequence 396, App  
27 37 56.1 73 9 US-10-176-488-396 Sequence 396, App  
28 37 56.1 73 9 US-10-176-492-396 Sequence 396, App  
29 37 56.1 73 9 US-10-176-747-396 Sequence 396, App  
30 37 56.1 73 9 US-10-176-750-396 Sequence 396, App  
31 37 56.1 73 9 US-10-176-985-396 Sequence 396, App  
32 37 56.1 73 9 US-10-176-987-396 Sequence 396, App  
33 37 56.1 73 9 US-10-176-991-396 Sequence 396, App  
34 37 56.1 73 9 US-10-176-992-396 Sequence 396, App  
35 37 56.1 73 9 US-10-176-993-396 Sequence 396, App  
36 37 56.1 73 9 US-10-184-658-396 Sequence 396, App  
37 37 56.1 73 9 US-10-173-695-396 Sequence 396, App  
38 37 56.1 73 9 US-10-173-697-396 Sequence 396, App  
39 37 56.1 73 9 US-10-173-705-396 Sequence 396, App  
40 37 56.1 73 9 US-10-174-576-396 Sequence 396, App  
41 37 56.1 73 9 US-10-174-585-396 Sequence 396, App  
42 37 56.1 73 9 US-10-174-586-396 Sequence 396, App  
43 37 56.1 73 9 US-10-175-747-396 Sequence 396, App  
44 37 56.1 73 9 US-10-176-481-396 Sequence 396, App  
45 37 56.1 73 9 US-10-176-485-396 Sequence 396, App

#### ALIGNMENTS

##### RESULT 1

US-09-934-251A-2  
; Sequence 2, Application US/09934251A  
; Patent No. US20020132017A1

; GENERAL INFORMATION:

; APPLICANT: Moore, Jeffrey G.

; TITLE OF INVENTION: Composition and Method for Preserving Progenitor Cells

; FILE REFERENCE: 108236.136

; CURRENT APPLICATION NUMBER: US/09/934,251A

; CURRENT FILING DATE: 2001-08-21

; PRIOR APPLICATION NUMBER: US 09/368,607

; PRIOR FILING DATE: 1999-08-05

; PRIOR APPLICATION NUMBER: US 08/762,537

; PRIOR FILING DATE: 1996-12-09

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 14

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: beta peptide sequence

US-09-934-251A-2

Query Match 98.5%; Score 65; DB 10; Length 14;

Best Local Similarity 92.9%; Pred. No. 3.2e-06;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFDLD 14

Db 1 AQSLSFSFTKFDLD 14

##### RESULT 2

US-09-934-251A-1

; Sequence 1, Application US/09934251A

; Patent No. US20020132017A1

; GENERAL INFORMATION:

; APPLICANT: Moore, Jeffrey G.

; TITLE OF INVENTION: Composition and Method for Preserving Progenitor Cells

; FILE REFERENCE: 108236.136

; CURRENT APPLICATION NUMBER: US/09/934,251A

; CURRENT FILING DATE: 2001-08-21  
 ; PRIOR APPLICATION NUMBER: US 09/368,607  
 ; PRIOR FILING DATE: 1999-08-05  
 ; PRIOR APPLICATION NUMBER: US 08/762,537  
 ; PRIOR FILING DATE: 1996-12-09  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: FASTSEQ for Windows Version 4.0  
 ; SEQ ID NO 1  
 ; LENGTH: 12  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: peptide chain of pylartin protein  
 US-09-934-251A-1

Query Match 83.38; Score 55; DB 10; Length 12;  
 Best Local Similarity 91.74; Pred. No. 0.0002;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDL 12  
 :|||:||||  
 Db 1 AQSLSFSFTKFD 12

RESULT 3  
 US-09-730-989-22  
 ; Sequence 22, Application US/09730989  
 ; Patent No. US20020061552A1  
 ; GENERAL INFORMATION:

; APPLICANT: Yan, Dong  
 ; APPLICANT: Williams, Lewis T.  
 ; TITLE OF INVENTION: MAMMALIAN DISHEVELLED-ASSOCIATED PROTEINS  
 ; FILE REFERENCE: PP-01657.002 / 200130.518  
 ; CURRENT APPLICATION NUMBER: US/09/730,989  
 ; CURRENT FILING DATE: 2001-05-08  
 ; NUMBER OF SEQ ID NOS: 29  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 22  
 ; LENGTH: 60  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila melanogaster  
 US-09-730-989-22

Query Match 63.6%; Score 42; DB 10; Length 60;  
 Best Local Similarity 57.1%; Pred. No. 0.26;  
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDL 14  
 :|||:||||  
 Db 18 SLPQLQSFYFDL 31

RESULT 4  
 US-09-730-989-20  
 ; Sequence 20, Application US/09730989  
 ; Patent No. US20020061552A1  
 ; GENERAL INFORMATION:

; APPLICANT: Yan, Dong  
 ; APPLICANT: Williams, Lewis T.  
 ; TITLE OF INVENTION: MAMMALIAN DISHEVELLED-ASSOCIATED PROTEINS  
 ; FILE REFERENCE: PP-01657.002 / 200130.518  
 ; CURRENT APPLICATION NUMBER: US/09/730,989  
 ; CURRENT FILING DATE: 2001-05-08  
 ; NUMBER OF SEQ ID NOS: 29  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 20  
 ; LENGTH: 668  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila melanogaster  
 US-09-730-989-20

Query Match 63.6%; Score 42; DB 10; Length 668;  
 Best Local Similarity 57.1%; Pred. No. 3;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 AQSLSFXFTKFDL 14  
 :|||:||||  
 Db 190 SLPQLQSFYFDL 203

## RESULT 5

US-10-063-547-148  
 ; Sequence 148, Application US/10063547  
 ; Publication No. US20020182638A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Eaton, Dan L.  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, Christopher J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Wood, William I.  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
 ; FILE REFERENCE: P3230R1C1  
 ; CURRENT APPLICATION NUMBER: US/10/063,547  
 ; CURRENT FILING DATE: 2002-05-02  
 ; Prior Application removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 170  
 ; SEQ ID NO 148  
 ; LENGTH: 73  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapien  
 US-10-063-547-148

Query Match 56.1%; Score 37; DB 9; Length 73;  
 Best Local Similarity 72.7%; Pred. No. 2.7;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDL 13  
 :|||:||||  
 Db 47 SLSFYFLKFQL 57

## RESULT 6

US-10-174-590-396  
 ; Sequence 396, Application US/10174590  
 ; Publication No. US20030008352A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Chen, Jian  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Wood, William I.  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
 ; FILE REFERENCE: P3430R1C42  
 ; CURRENT APPLICATION NUMBER: US/10/174,590  
 ; CURRENT FILING DATE: 2002-06-18  
 ; Prior application removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 612  
 ; SEQ ID NO 396  
 ; LENGTH: 73  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapien  
 US-10-174-590-396

Query Match 56.1%; Score 37; DB 9; Length 73;

Best Local Similarity 72.7%; Pred. No. 2.7;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSEFXFTKFDL 13  
|||||  
Db 47 SLSEFYFLKFQL 57

## RESULT 7

US-10-176-758-396  
; Sequence 396, Application US/10176758  
; Publication No. US20030008353A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3430R1C104  
; CURRENT APPLICATION NUMBER: US/10/176,758  
; CURRENT FILING DATE: 2002-06-21  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 396  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-176-758-396

Query Match 56.1%; Score 37; DB 9; Length 73;  
Best Local Similarity 72.7%; Pred. No. 2.7;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSEFXFTKFDL 13  
|||||  
Db 47 SLSEFYFLKFQL 57

## RESULT 8

US-10-063-616-148  
; Sequence 148, Application US/10063616  
; Publication No. US20030013855A1  
; GENERAL INFORMATION:  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3230R1C1  
; CURRENT APPLICATION NUMBER: US/10/063,616  
; CURRENT FILING DATE: 2002-05-03  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 170  
; SEQ ID NO 148  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-063-616-148

Query Match 56.1%; Score 37; DB 9; Length 73;  
Best Local Similarity 72.7%; Pred. No. 2.7;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSEFXFTKFDL 13  
|||||  
Db 47 SLSEFYFLKFQL 57

## RESULT 9

US-10-175-737-396  
; Sequence 396, Application US/10175737  
; Publication No. US20030013153A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3430R1C150  
; CURRENT APPLICATION NUMBER: US/10/175,737  
; CURRENT FILING DATE: 2002-06-19  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 396  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-175-737-396

Query Match 56.1%; Score 37; DB 9; Length 73;  
Best Local Similarity 72.7%; Pred. No. 2.7;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSEFXFTKFDL 13  
|||||  
Db 47 SLSEFYFLKFQL 57

## RESULT 10

US-10-063-502-148  
; Sequence 148, Application US/10063502  
; Publication No. US20030023042A1  
; GENERAL INFORMATION:  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3230R1C1  
; CURRENT APPLICATION NUMBER: US/10/063,502  
; CURRENT FILING DATE: 2002-05-01  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 170  
; SEQ ID NO 148  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-063-502-148

Query Match 56.1%; Score 37; DB 9; Length 73;  
Best Local Similarity 72.7%; Pred. No. 2.7;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDL 13  
|||||  
Db 47 SLSFYFLKQQL 57

## RESULT 11

US-10-173-706-396  
; Sequence 396, Application US/10173706  
; Publication No. US2003002293A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C7  
; CURRENT APPLICATION NUMBER: US/10/173,706  
; CURRENT FILING DATE: 2002-06-17  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 396  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-173-706-396

Query Match 56.1%; Score 37; DB 9; Length 73;  
Best Local Similarity 72.7%; Pred. No. 2.7;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDL 13  
|||||  
Db 47 SLSFYFLKQQL 57

## RESULT 12

US-10-175-738-396  
; Sequence 396, Application US/10175738  
; Publication No. US2003002294A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C45  
; CURRENT APPLICATION NUMBER: US/10/175,738  
; CURRENT FILING DATE: 2002-06-19  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 396  
; LENGTH: 73

; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-175-738-396

Query Match 56.1%; Score 37; DB 9; Length 73;  
Best Local Similarity 72.7%; Pred. No. 2.7;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDL 13  
|||||  
Db 47 SLSFYFLKQQL 57

## RESULT 13

US-10-175-752-396  
; Sequence 396, Application US/10175752  
; Publication No. US2003002295A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C60  
; CURRENT APPLICATION NUMBER: US/10/175,752  
; CURRENT FILING DATE: 2002-06-19  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 396  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-175-752-396

Query Match 56.1%; Score 37; DB 9; Length 73;  
Best Local Similarity 72.7%; Pred. No. 2.7;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDL 13  
|||||  
Db 47 SLSFYFLKQQL 57

## RESULT 14

US-10-176-482-396  
; Sequence 396, Application US/10176482  
; Publication No. US2003002296A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C70  
; CURRENT APPLICATION NUMBER: US/10/176,482  
; CURRENT FILING DATE: 2002-06-20  
; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 396  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-176-482-396

Query Match 56.1%; Score 37; DB 9; Length 73;  
Best Local Similarity 72.7%; Pred. No. 2.7;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDL 13  
| | | | | | | |  
Db 47 SLSFYFLKFQL 57

## RESULT 15

US-10-176-757-396

; Sequence 396, Application US/10176757

; Publication No. US20030022297A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3430R1C86

; CURRENT APPLICATION NUMBER: US/10/176,757

; CURRENT FILING DATE: 2002-06-20

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 396

; LENGTH: 73

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-176-757-396

Query Match 56.1%; Score 37; DB 9; Length 73;  
Best Local Similarity 72.7%; Pred. No. 2.7;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDL 13  
| | | | | | | |  
Db 47 SLSFYFLKFQL 57

Search completed: February 26, 2003, 14:55:42  
Job time : 13 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:49:06 : Search time 139 Seconds  
(without alignments)  
64.937 Million cell updates/sec

Title: US-09-476-485A-31  
Perfect score: 66  
Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main.\*

|     |                                        |
|-----|----------------------------------------|
| 1:  | /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.* |
| 2:  | /cgn2_6/ptodata/1/paa/US06_COMB.pep.*  |
| 3:  | /cgn2_6/ptodata/1/paa/US07_COMB.pep.*  |
| 4:  | /cgn2_6/ptodata/1/paa/US08_COMB.pep.*  |
| 5:  | /cgn2_6/ptodata/1/paa/US081_COMB.pep.* |
| 6:  | /cgn2_6/ptodata/1/paa/US082_COMB.pep.* |
| 7:  | /cgn2_6/ptodata/1/paa/US083_COMB.pep.* |
| 8:  | /cgn2_6/ptodata/1/paa/US084_COMB.pep.* |
| 9:  | /cgn2_6/ptodata/1/paa/US085_COMB.pep.* |
| 10: | /cgn2_6/ptodata/1/paa/US086_COMB.pep.* |
| 11: | /cgn2_6/ptodata/1/paa/US087_COMB.pep.* |
| 12: | /cgn2_6/ptodata/1/paa/US088_COMB.pep.* |
| 13: | /cgn2_6/ptodata/1/paa/US089_COMB.pep.* |
| 14: | /cgn2_6/ptodata/1/paa/US090_COMB.pep.* |
| 15: | /cgn2_6/ptodata/1/paa/US091_COMB.pep.* |
| 16: | /cgn2_6/ptodata/1/paa/US092_COMB.pep.* |
| 17: | /cgn2_6/ptodata/1/paa/US093_COMB.pep.* |
| 18: | /cgn2_6/ptodata/1/paa/US094_COMB.pep.* |
| 19: | /cgn2_6/ptodata/1/paa/US095_COMB.pep.* |
| 20: | /cgn2_6/ptodata/1/paa/US096_COMB.pep.* |
| 21: | /cgn2_6/ptodata/1/paa/US097_COMB.pep.* |
| 22: | /cgn2_6/ptodata/1/paa/US098_COMB.pep.* |
| 23: | /cgn2_6/ptodata/1/paa/US099_COMB.pep.* |
| 24: | /cgn2_6/ptodata/1/paa/US100_COMB.pep.* |
| 25: | /cgn2_6/ptodata/1/paa/US101_COMB.pep.* |
| 26: | /cgn2_6/ptodata/1/paa/US102_COMB.pep.* |
| 27: | /cgn2_6/ptodata/1/paa/US60_COMB.pep.*  |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 65    | 98.5        | 14     | 1  | PCT-US97-22486-2  |
| 2          | 65    | 98.5        | 14     | 23 | US-09-934-251A-2  |
| 3          | 64    | 97.0        | 14     | 4  | US-08-081-508-12  |
| 4          | 64    | 97.0        | 14     | 5  | US-08-157-490-11  |
| 5          | 64    | 97.0        | 14     | 18 | US-09-476-485A-31 |
| 6          | 64    | 97.0        | 279    | 18 | US-09-476-485A-56 |

|    |      |      |     |    |                       |                    |
|----|------|------|-----|----|-----------------------|--------------------|
| 7  | 64   | 97.0 | 279 | 21 | US-09-791-537-109055  | Sequence 109055,   |
| 8  | 64   | 97.0 | 303 | 18 | US-09-476-485A-6      | Sequence 6, Appli  |
| 9  | 56   | 84.8 | 15  | 18 | US-09-476-485A-9      | Sequence 9, Appli  |
| 10 | 56   | 84.8 | 234 | 18 | US-09-476-485A-8      | Sequence 8, Appli  |
| 11 | 55   | 83.3 | 12  | 1  | PCT-US97-22486-1      | Sequence 1, Appli  |
| 12 | 55   | 83.3 | 12  | 23 | US-09-934-251A-1      | Sequence 1, Appli  |
| 13 | 55   | 83.3 | 105 | 1  | PCT-US98-13046-13     | Sequence 13, Appli |
| 14 | 55   | 83.3 | 105 | 24 | US-10-045-353-13      | Sequence 13, Appli |
| 15 | 55   | 83.3 | 123 | 18 | US-09-476-485A-50     | Sequence 50, Appli |
| 16 | 55   | 83.3 | 237 | 18 | US-09-476-485A-55     | Sequence 55, Appli |
| 17 | 55   | 83.3 | 270 | 1  | PCT-US98-13046-2      | Sequence 2, Appli  |
| 18 | 55   | 83.3 | 270 | 24 | US-10-045-353-2       | Sequence 2, Appli  |
| 19 | 55   | 83.3 | 272 | 21 | US-09-791-537-1173    | Sequence 1173, Ap  |
| 20 | 55   | 83.3 | 286 | 1  | PCT-US98-13046-23     | Sequence 23, Appli |
| 21 | 55   | 83.3 | 286 | 18 | US-09-476-485A-23     | Sequence 23, Appli |
| 22 | 55   | 83.3 | 286 | 24 | US-10-045-353-23      | Sequence 23, Appli |
| 23 | 48   | 72.7 | 264 | 18 | US-09-476-485A-2      | Sequence 2, Appli  |
| 24 | 47   | 71.2 | 14  | 18 | US-09-476-485A-32     | Sequence 32, Appli |
| 25 | 44.5 | 67.4 | 106 | 18 | US-09-476-485A-49     | Sequence 49, Appli |
| 26 | 42   | 63.6 | 60  | 21 | US-09-730-989-22      | Sequence 22, Appli |
| 27 | 42   | 63.6 | 195 | 16 | US-09-270-767-61659   | Sequence 61659, A  |
| 28 | 42   | 63.6 | 195 | 16 | US-09-270-8498-193357 | Sequence 193357, A |
| 29 | 42   | 63.6 | 280 | 16 | US-09-252-991A-16878  | Sequence 16878, A  |
| 30 | 42   | 63.6 | 668 | 21 | US-09-730-989-20      | Sequence 20, Appli |
| 31 | 42   | 63.6 | 687 | 16 | US-09-270-767-46104   | Sequence 46104, A  |
| 32 | 42   | 63.6 | 895 | 27 | US-60-173-464-22125   | Sequence 22125, A  |
| 33 | 42   | 63.6 | 896 | 20 | US-09-614-150-33690   | Sequence 33690, A  |
| 34 | 42   | 63.6 | 928 | 1  | PCT-US00-04188-2      | Sequence 2, Appli  |
| 35 | 42   | 63.6 | 928 | 19 | US-09-506-066-2       | Sequence 2, Appli  |
| 36 | 41   | 62.1 | 44  | 19 | US-09-513-996A-72209  | Sequence 72209, A  |
| 37 | 41   | 62.1 | 44  | 19 | US-09-513-996A-78778  | Sequence 78778, A  |
| 38 | 40   | 60.6 | 373 | 25 | US-10-179-131-7061    | Sequence 7061, Ap  |
| 39 | 39   | 59.1 | 279 | 21 | US-09-791-537-115612  | Sequence 115612, A |
| 40 | 39   | 59.1 | 968 | 16 | US-09-252-991A-17402  | Sequence 17402, A  |
| 41 | 38.5 | 58.3 | 13  | 18 | US-09-476-485A-34     | Sequence 34, Appli |
| 42 | 38   | 57.6 | 254 | 21 | US-09-791-537-112588  | Sequence 112588, A |
| 43 | 38   | 57.6 | 458 | 21 | US-09-733-089-18926   | Sequence 18926, A  |
| 44 | 38   | 57.6 | 458 | 21 | US-09-733-089-19024   | Sequence 19024, A  |
| 45 | 38   | 57.6 | 458 | 22 | US-09-816-660-18926   | Sequence 18926, A  |

## ALIGNMENTS

RESULT 1  
PCT-US97-22486-2  
; Sequence 2, Application PC/TUS9722486  
; GENERAL INFORMATION:  
; APPLICANT: Moore, Jeffrey G.  
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR  
; NUMBER OF INVENTIONS: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron, LLP  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11753  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US97/22486  
; FILING DATE: 9-DEC-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/762,537  
; FILING DATE: 9-DEC-1996  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: 08/825,369
; FILING DATE: 28-MAR-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Dea, Sean W. 37690
; REGISTRATION NUMBER: 381-21 CIP/PCT
; REFERENCE/DOCKET NUMBER: 381-21 CIP/PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US97-22486-2

```

```

Query Match      98.5%; Score 65; DB 1; Length 14;
Best Local Similarity 92.9%; Pred. No. 0.0001;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 AQSLSFXFTKFDLD 14
   ||||| |||||
DB 1 AQSLSFSFTKFDLD 14

```

```

RESULT 2
US-09-934-251A-2
; Sequence 2, Application US/09934251A
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Composition and Method for Preserving Progenitor Cells
; FILE REFERENCE: 108236.136
; CURRENT APPLICATION NUMBER: US/09/934,251A
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 09/368,607
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 08/762,537
; PRIOR FILING DATE: 1996-12-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 14
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: beta peptide sequence
US-09-934-251A-2

```

```

Query Match      98.5%; Score 65; DB 23; Length 14;
Best Local Similarity 92.9%; Pred. No. 0.0001;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 AQSLSFXFTKFDLD 14
   ||||| |||||
DB 1 AQSLSFSFTKFDLD 14

```

```

RESULT 3
US-08-081-508-12
; Sequence 12, Application US/08081508
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TROPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.

```

```

; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/081,508
; FILING DATE: 19930621
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/975,049
; FILING DATE: 12-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,451
; FILING DATE: 19-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/005,941
; FILING DATE: 15-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/045,272
; FILING DATE: 01-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US UNASSIGNED
; FILING DATE: 09-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US UNASSIGNED
; FILING DATE: 18-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N. 28,601
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-12P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-081-508-12

```

```

Query Match      97.0%; Score 64; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 AQSLSFXFTKFDLD 14
   ||||| |||||
DB 1 AQSLSFXFTKFDLD 14

```

```

RESULT 4
US-08-157-490-11

```



Sequence 11, Application US/08157490  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
TOPLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
US-08-157-490-11

LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
US-08-157-490-11

Query Match 97.0%; Score 64; DB 5; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFDLD 14  
Db 1 AQSLSFXFTKFDLD 14

RESULT 5  
US-09-476-485A-31  
Sequence 31, Application US/09476485A  
GENERAL INFORMATION:  
APPLICANT: Colucci, M. Gabriella  
APPLICANT: Moore, Jeffrey G.  
TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for  
TITLE OF INVENTION: and products of Their Use  
FILE REFERENCE: 108236.119  
CURRENT APPLICATION NUMBER: US/09/476,485A  
CURRENT FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: US 08/881,189  
PRIOR FILING DATE: 1997-06-24  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 31  
LENGTH: 14  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Peptide corresponding to Pv-FRIL.  
NAME/KEY: PEPTIDE  
LOCATION: (7)..(7)  
OTHER INFORMATION: Amino acid 7 is Xaa wherein Xaa = Asn, Cys or Ser.  
US-09-476-485A-31

Query Match 97.0%; Score 64; DB 18; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFDLD 14  
Db 1 AQSLSFXFTKFDLD 14

RESULT 6  
US-09-476-485A-56  
Sequence 56, Application US/09476485A  
GENERAL INFORMATION:  
APPLICANT: Colucci, M. Gabriella  
APPLICANT: Chrispeels, Maarten J.  
APPLICANT: Moore, Jeffrey G.  
TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for  
TITLE OF INVENTION: and products of Their Use  
FILE REFERENCE: 108236.119  
CURRENT APPLICATION NUMBER: US/09/476,485A  
CURRENT FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: US 08/881,189  
PRIOR FILING DATE: 1997-06-24  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 56

Sequence 11, Application US/08157490  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
TOPLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
US-08-157-490-11

LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
US-08-157-490-11

Query Match 97.0%; Score 64; DB 5; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFDLD 14  
Db 1 AQSLSFXFTKFDLD 14

RESULT 5  
US-09-476-485A-31  
Sequence 31, Application US/09476485A  
GENERAL INFORMATION:  
APPLICANT: Colucci, M. Gabriella  
APPLICANT: Moore, Jeffrey G.  
TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for  
TITLE OF INVENTION: and products of Their Use  
FILE REFERENCE: 108236.119  
CURRENT APPLICATION NUMBER: US/09/476,485A  
CURRENT FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: US 08/881,189  
PRIOR FILING DATE: 1997-06-24  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 31  
LENGTH: 14  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Peptide corresponding to Pv-FRIL.  
NAME/KEY: PEPTIDE  
LOCATION: (7)..(7)  
OTHER INFORMATION: Amino acid 7 is Xaa wherein Xaa = Asn, Cys or Ser.  
US-09-476-485A-31

Query Match 97.0%; Score 64; DB 18; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFDLD 14  
Db 1 AQSLSFXFTKFDLD 14

RESULT 6  
US-09-476-485A-56  
Sequence 56, Application US/09476485A  
GENERAL INFORMATION:  
APPLICANT: Colucci, M. Gabriella  
APPLICANT: Chrispeels, Maarten J.  
APPLICANT: Moore, Jeffrey G.  
TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for  
TITLE OF INVENTION: and products of Their Use  
FILE REFERENCE: 108236.119  
CURRENT APPLICATION NUMBER: US/09/476,485A  
CURRENT FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: US 08/881,189  
PRIOR FILING DATE: 1997-06-24  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 56

; LENGTH: 279  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: PvFRIL.  
US-09-476-485A-56

Query Match 97.0%; Score 64; DB 18; Length 279;  
Best Local Similarity 92.9%; Pred. No. 0.0034;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14  
||||| |||||  
Db 1 AQSLSFNFTKFDLD 14

RESULT 7  
US-09-791-537-109055  
; Sequence 109055, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 109055  
; LENGTH: 279  
; TYPE: PRT  
; ORGANISM: Phaseolus vulgaris  
US-09-791-537-109055

Query Match 97.0%; Score 64; DB 21; Length 279;  
Best Local Similarity 92.9%; Pred. No. 0.0034;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14  
||||| |||||  
Db 1 AQSLSFNFTKFDLD 14

RESULT 8  
US-09-476-485A-6  
; Sequence 6, Application US/09476485A  
; GENERAL INFORMATION:  
; APPLICANT: Colucci, M. Gabriella  
; APPLICANT: Chrispeels, Maarten J.  
; APPLICANT: Moore, Jeffrey G.  
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for  
; TITLE OF INVENTION: and Products of Their Use  
; FILE REFERENCE: 108236.119  
; CURRENT APPLICATION NUMBER: US/09/476,485A  
; CURRENT FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 08/881,189  
; PRIOR FILING DATE: 1997-06-24  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 6  
; LENGTH: 303  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Pv-FRIL.  
US-09-476-485A-6

Query Match 97.0%; Score 64; DB 18; Length 303;  
Best Local Similarity 92.9%; Pred. No. 0.0037;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14  
||||| |||||  
Db 1 AQSLSFNFTKFDLD 14

RESULT 9  
US-09-476-485A-9  
; Sequence 9, Application US/09476485A  
; GENERAL INFORMATION:  
; APPLICANT: Colucci, M. Gabriella  
; APPLICANT: Chrispeels, Maarten J.  
; APPLICANT: Moore, Jeffrey G.  
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for  
; TITLE OF INVENTION: and Products of Their Use  
; FILE REFERENCE: 108236.119  
; CURRENT APPLICATION NUMBER: US/09/476,485A  
; CURRENT FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 08/881,189  
; PRIOR FILING DATE: 1997-06-24  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 9  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Beta band polypeptide.  
US-09-476-485A-9

Query Match 84.8%; Score 56; DB 18; Length 15;  
Best Local Similarity 78.6%; Pred. No. 0.0051;  
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14  
||||| |||||  
Db 1 AQSLSFNFTKFDSD 14

RESULT 10  
US-09-476-485A-8  
; Sequence 8, Application US/09476485A  
; GENERAL INFORMATION:  
; APPLICANT: Colucci, M. Gabriella  
; APPLICANT: Chrispeels, Maarten J.  
; APPLICANT: Moore, Jeffrey G.  
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for  
; TITLE OF INVENTION: and Products of Their Use  
; FILE REFERENCE: 108236.119  
; CURRENT APPLICATION NUMBER: US/09/476,485A  
; CURRENT FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 08/881,189  
; PRIOR FILING DATE: 1997-06-24  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 8  
; LENGTH: 234  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: YamPril deduced amino acid sequence.  
US-09-476-485A-8

Query Match 84.8%; Score 56; DB 18; Length 234;  
Best Local Similarity 78.6%; Pred. No. 0.086; 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14  
||||| |||||  
Db 1 AQSLSFNFTKFDSD 14

RESULT 11  
PCT-US97-22486-1

; Sequence 1, Application PC/TUS9722486  
; GENERAL INFORMATION:  
; APPLICANT: Moore, Jeffrey G.  
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR  
; PRESERVING PROGENITOR CELLS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron, LLP  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11753  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US97/22486  
; FILING DATE: 9-DEC-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/762,537  
; FILING DATE: 9-DEC-1996  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/825,369  
; FILING DATE: 28-MAR-1997  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: O'Dea, Sean W.  
; REGISTRATION NUMBER: 37690  
; REFERENCE/DOCKET NUMBER: 381-21 CIP/PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US97-22486-1

Query Match 83.3%; Score 55; DB 1; Length 12;  
Best Local Similarity 91.7%; Pred. No. 0.0063;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 AQSLSFXFTKFD 12  
Db 1 AQSLSFSFTKFD 12

RESULT 12  
US-09-934-251A-1  
; Sequence 1, Application US/09934251A  
; GENERAL INFORMATION:  
; APPLICANT: Moore, Jeffrey G.  
; TITLE OF INVENTION: Composition and Method for Preserving Progenitor Cells  
; FILE REFERENCE: 108236.136  
; CURRENT APPLICATION NUMBER: US/09/934, 251A  
; CURRENT FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 09/368,607  
; PRIOR FILING DATE: 1999-08-05  
; PRIOR APPLICATION NUMBER: US 08/762,537  
; PRIOR FILING DATE: 1996-12-09  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Artificial Sequence

; FEATURE:  
; OTHER INFORMATION: peptide chain of pylartin protein  
US-09-934-251A-1  
Query Match 83.3%; Score 55; DB 23; Length 12;  
Best Local Similarity 91.7%; Pred. No. 0.0063;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 AQSLSFXFTKFD 12  
Db 1 AQSLSFSFTKFD 12  
RESULT 13  
PCT-US98-13046-13  
; Sequence 13, Application PC/TUS9813046  
; GENERAL INFORMATION:  
; APPLICANT: Colucci et al.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron, LLP  
; STREET: 6900 Jericho Turnpike  
; CITY: Syosset  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11791  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US98/13046  
; FILING DATE: June 23, 1998  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/881,189  
; FILING DATE: June 24, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: 381-44 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 105 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US98-13046-13

Query Match 83.3%; Score 55; DB 1; Length 105;  
Best Local Similarity 91.7%; Pred. No. 0.058;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 AQSLSFXFTKFD 12  
Db 1 AQSLSFSFTKFD 12

RESULT 14  
US-10-045-353-13  
; Sequence 13, Application US/10045353  
; GENERAL INFORMATION:  
; APPLICANT: Colucci et al.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED  
; PROGENITOR CELL PRESERVATION FACTOR  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron, LLP

STREET: 350 Jericho Turnpike  
CITY: Jericho  
STATE: New York  
COUNTRY: USA  
ZIP: 11753

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
COMPUTER: IBM compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/045,353  
FILING DATE: 29-Oct-2001  
CLASSIFICATION: <Unknown>

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/881,189

## FILING DATE: &lt;Unknown&gt;

## ATTORNEY/AGENT INFORMATION:

NAME: Feit, Irving N.

REGISTRATION NUMBER: 28,601

REFERENCE/DOCKET NUMBER: 381-44 PCT

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 822-3550

TELEFAX: (516) 822-5582

## INFORMATION FOR SEQ ID NO: 13:

## SEQUENCE CHARACTERISTICS:

LENGTH: 105 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 13:

US-10-045-353-13

Query Match 83.3%; Score 55; DB 24; Length 105;  
Best Local Similarity 91.7%; Pred. No. 0.058;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12  
Db 1 AQSLSFSFTKFD 12

## RESULT 15

US-09-476-485A-50

; Sequence 50, Application US/09476485A

; GENERAL INFORMATION:

; APPLICANT: Colucci, M. Gabriella

; APPLICANT: Chrispeels, Maarten J.

; APPLICANT: Moore, Jeffrey G.

; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for

; FILE REFERENCE: 108236.119

; CURRENT APPLICATION NUMBER: US/09/476,485A

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: US 08/881,189

; PRIOR FILING DATE: 1997-06-24

; NUMBER OF SEQ ID NOS: 57

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 50

; LENGTH: 123

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Beta-subunit of D1-FRIL.

US-09-476-485A-50

Query Match 83.3%; Score 55; DB 18; Length 123;  
Best Local Similarity 91.7%; Pred. No. 0.068;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12  
Db 1 AQSLSFSFTKFD 12

Search completed: February 26, 2003, 14:54:16  
Job time : 141 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:49:30 ; Search time 60 Seconds  
(without alignments)  
21.326 Million cell updates/sec

Title: US-09-476-485A-31

Perfect score: 66

Sequence: 1 AQSLSFXFKFDLD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 486122 seqs, 91396495 residues

Total number of hits satisfying chosen parameters: 486122

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*

- 1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description          |
|------------|-------|-------------|--------|-------|----------------------|
| 1          | 64    | 97.0        | 279    | 6     | US-10-190-258A-6     |
| 2          | 64    | 97.0        | 303    | 6     | US-10-083-936B-6     |
| 3          | 56    | 84.8        | 15     | 6     | US-10-083-936B-9     |
| 4          | 56    | 84.8        | 15     | 6     | US-10-190-258A-9     |
| 5          | 56    | 84.8        | 234    | 6     | US-10-083-936B-8     |
| 6          | 56    | 84.8        | 234    | 6     | US-10-190-258A-8     |
| 7          | 55    | 83.3        | 264    | 6     | US-10-083-936B-2     |
| 8          | 55    | 83.3        | 264    | 6     | US-10-190-258A-2     |
| 9          | 37    | 56.1        | 73     | 6     | US-10-125-923A-396   |
| 10         | 37    | 56.1        | 73     | 6     | US-10-063-580-148    |
| 11         | 37    | 56.1        | 73     | 6     | US-10-205-892-396    |
| 12         | 37    | 56.1        | 73     | 6     | US-10-174-575-396    |
| 13         | 37    | 56.1        | 73     | 6     | US-10-174-575A-396   |
| 14         | 37    | 56.1        | 73     | 6     | US-10-187-755-396    |
| 15         | 37    | 56.1        | 73     | 6     | US-10-187-749-396    |
| 16         | 37    | 56.1        | 73     | 6     | US-10-199-672-396    |
| 17         | 37    | 56.1        | 73     | 6     | US-10-063-557-148    |
| 18         | 37    | 56.1        | 73     | 6     | US-10-194-486-396    |
| 19         | 37    | 56.1        | 1108   | 5     | US-09-958-109A-6     |
| 20         | 35    | 53.0        | 139    | 6     | US-10-203-138A-11864 |
| 21         | 35    | 53.0        | 168    | 5     | US-09-724-676-71898  |
| 22         | 35    | 53.0        | 168    | 5     | US-09-724-676-71900  |
| 23         | 35    | 53.0        | 168    | 5     | US-09-724-676-71903  |
| 24         | 35    | 53.0        | 168    | 5     | US-09-724-676-71904  |
| 25         | 35    | 53.0        | 168    | 5     | US-09-724-676A-71898 |
| 26         | 35    | 53.0        | 168    | 5     | US-09-724-676A-71900 |

27 35 53.0 168 5 US-09-724-676A-71903 Sequence 71903, A  
28 35 53.0 168 5 US-09-724-676A-71904 Sequence 71904, A  
29 35 53.0 191 5 US-09-724-676-71899 Sequence 71899, A  
30 35 53.0 191 5 US-09-724-676-71902 Sequence 71902, A  
31 35 53.0 191 5 US-09-724-676A-71899 Sequence 71899, A  
32 35 53.0 191 5 US-09-724-676A-71902 Sequence 71902, A  
33 35 53.0 193 1 PCT-US02-29560-268 Sequence 268, App  
34 35 53.0 193 1 PCT-US02-29560-269 Sequence 269, App  
35 35 53.0 193 5 US-09-724-676-48958 Sequence 48958, A  
36 35 53.0 193 5 US-09-724-676-48959 Sequence 48959, A  
37 35 53.0 193 5 US-09-724-676A-48958 Sequence 48958, A  
38 35 53.0 193 5 US-09-724-676A-48959 Sequence 48959, A  
39 35 53.0 193 6 US-10-245-882-268 Sequence 268, App  
40 35 53.0 193 6 US-10-245-882-269 Sequence 269, App  
41 35 53.0 329 6 US-10-050-902-217 Sequence 217, App  
42 35 53.0 329 6 US-10-346-190-11 Sequence 11, Appl  
43 35 53.0 329 6 US-10-050-898-217 Sequence 217, Appl  
44 35 53.0 329 6 US-10-289-454-11 Sequence 11, Appl  
45 35 53.0 625 6 US-10-092-411A-4504 Sequence 4504, Ap

## ALIGNMENTS

RESULT 1  
US-10-190-258A-6  
; Sequence 6, Application US/10190258A  
; GENERAL INFORMATION:  
; APPLICANT: Moore, Jeffrey G  
; TITLE OF INVENTION: Dendritic Cell Isolation Methods  
; FILE REFERENCE: 108236.132  
; CURRENT APPLICATION NUMBER: US/10/190,258A  
; CURRENT FILING DATE: 2002-07-03  
; PRIOR APPLICATION NUMBER: US 60/303,265  
; PRIOR FILING DATE: 2001-07-05  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 279  
; TYPE: PRT  
; ORGANISM: Phaseolus vulgaris  
US-10-190-258A-6

Query Match 97.0%; Score 64; DB 6; Length 279;  
Best Local Similarity 92.9%; Pred. No. 2.1e-05;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFKFDLD 14  
| | | | | | | | | |  
Db 1 AQSLSFNFTKFDLD 14

RESULT 2  
US-10-083-936B-6  
; Sequence 6, Application US/10083936B  
; GENERAL INFORMATION:  
; APPLICANT: Moore, Jeffrey G.  
; TITLE OF INVENTION: Compositions and Methods for Protecting Tissues and  
; FILE REFERENCE: 108236.130  
; CURRENT APPLICATION NUMBER: US/10/083,936B  
; CURRENT FILING DATE: 2002-10-22  
; PRIOR APPLICATION NUMBER: US 60/271,666  
; PRIOR FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: US 60/302,716  
; PRIOR FILING DATE: 2001-07-03  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 303  
; TYPE: PRT  
; ORGANISM: Phaseolus vulgaris  
US-10-083-936B-6



; TYPE: PRT  
; ORGANISM: Dolichos lablab  
US-10-083-936B-2

Query Match 83.3%; Score 55; DB 6; Length 264;  
Best Local Similarity 91.7%; Pred. No. 0.0014;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12  
|||||  
Db 1 AQSLSFSFTKFD 12

## RESULT 8

US-10-190-258A-2

; Sequence 2, Application US/10190258A

; GENERAL INFORMATION:

; APPLICANT: Moore, Jeffrey G

; TITLE OF INVENTION: Dendritic Cell Isolation Methods

; FILE REFERENCE: 108236.132

; CURRENT APPLICATION NUMBER: US/10/190,258A

; CURRENT FILING DATE: 2002-07-03

; PRIOR APPLICATION NUMBER: US 60/303,265

; PRIOR FILING DATE: 2001-07-05

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 264

; TYPE: PRT

; ORGANISM: Dolichos lablab

US-10-190-258A-2

## Query Match

Best Local Similarity 83.3%; Score 55; DB 6; Length 264;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12  
|||||  
Db 1 AQSLSFSFTKFD 12

## RESULT 9

US-10-125-923A-396

; Sequence 396, Application US/10125923A

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C79

; CURRENT APPLICATION NUMBER: US/10/125,923A

; CURRENT FILING DATE: 2002-01-15

; PRIOR APPLICATION NUMBER: 10/052586

; PRIOR FILING DATE: 2002-01-15

; PRIOR APPLICATION NUMBER: 60/059263

; PRIOR FILING DATE: 1997-09-18

; PRIOR APPLICATION NUMBER: 60/059266

; PRIOR FILING DATE: 1997-09-18

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/063120

; PRIOR FILING DATE: 1997-10-24

; PRIOR APPLICATION NUMBER: 60/063121

; PRIOR FILING DATE: 1997-10-24

; PRIOR APPLICATION NUMBER: 60/063486  
; PRIOR FILING DATE: 1997-10-21  
; PRIOR APPLICATION NUMBER: 60/063540  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063541  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063544  
; PRIOR FILING DATE: 1997-10-28  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 396  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-125-923A-396

Query Match 56.1%; Score 37; DB 6; Length 73;  
Best Local Similarity 72.7%; Pred. No. 1.5;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDL 13  
|||||  
Db 47 SLSFYFLKFL 57

## RESULT 10

US-10-063-580-148

; Sequence 148, Application US/10063580

; GENERAL INFORMATION:

; APPLICANT: Eaton, Dan L.

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gerlitsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3230R1C1

; CURRENT APPLICATION NUMBER: US/10/063,580

; CURRENT FILING DATE: 2002-05-03

; PRIOR APPLICATION NUMBER: 60/063435

; PRIOR FILING DATE: 1997-10-29

; PRIOR APPLICATION NUMBER: 60/064215

; PRIOR FILING DATE: 1997-10-29

; PRIOR APPLICATION NUMBER: 60/082797

; PRIOR FILING DATE: 1998-04-22

; PRIOR APPLICATION NUMBER: 60/083495

; PRIOR FILING DATE: 1998-04-29

; PRIOR APPLICATION NUMBER: 60/085579

; PRIOR FILING DATE: 1998-05-15

; PRIOR APPLICATION NUMBER: 60/087759

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/088021

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088029

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088030

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088734

; PRIOR FILING DATE: 1998-06-10

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 170

; SEQ ID NO 148

; LENGTH: 73

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-063-580-148

## Query Match

Best Local Similarity 56.1%; Score 37; DB 6; Length 73;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 SLSFXFTKFDL 13  
||||| | | | |

Db 47 SLSFYFLKFQL 57

## RESULT 11

US-10-205-892-396  
; Sequence 396, Application US/10205892  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3430R1C35  
; CURRENT APPLICATION NUMBER: 2002-07-26  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: 10/052586  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059266  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063120  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063121  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063486  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063540  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063541  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063544  
; PRIOR FILING DATE: 1997-10-28  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 396  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Homo Sapien

US-10-205-892-396

Query Match 56.1%; Score 37; DB 6; Length 73;

Best Local Similarity 72.7%; Pred. No. 1.5;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 SLSFXFTKFDL 13  
||||| | | | |

Db 47 SLSFYFLKFQL 57

## RESULT 12

US-10-174-575-396  
; Sequence 396, Application US/10174575  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3430R1C35  
; CURRENT APPLICATION NUMBER: 2002-06-18  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: 10/052586  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059266  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063120  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063121  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063486  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063540  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063541  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063544  
; PRIOR FILING DATE: 1997-10-28  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 396  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Homo Sapien

US-10-174-575-396

APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
TITLE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P3430R1C35  
CURRENT APPLICATION NUMBER: US/10174,575  
CURRENT FILING DATE: 2002-06-18  
PRIOR APPLICATION NUMBER: 10/052586  
PRIOR FILING DATE: 2002-01-15  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059266  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/063120  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063121  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063486  
PRIOR FILING DATE: 1997-10-21  
PRIOR APPLICATION NUMBER: 60/063540  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063541  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063544  
PRIOR FILING DATE: 1997-10-28  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 396  
LENGTH: 73  
TYPE: PRT  
ORGANISM: Homo Sapien

US-10-174-575A-396

Query Match 56.1%; Score 37; DB 6; Length 73;

Best Local Similarity 72.7%; Pred. No. 1.5;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 SLSFXFTKFDL 13  
||||| | | | |

Db 47 SLSFYFLKFQL 57

## RESULT 13

US-10-174-575A-396  
; Sequence 396, Application US/10174575A  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3430R1C35  
; CURRENT APPLICATION NUMBER: US/10174,575A  
; CURRENT FILING DATE: 2002-06-18  
; PRIOR APPLICATION NUMBER: 10/052586  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059266  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063120  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063121  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063486  
; PRIOR FILING DATE: 1997-10-21  
; PRIOR APPLICATION NUMBER: 60/063540  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063541  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063544  
; PRIOR FILING DATE: 1997-10-28  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 396  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Homo Sapien

US-10-174-575A-396



; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063120  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063121  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063486  
; PRIOR FILING DATE: 1997-10-21  
; PRIOR APPLICATION NUMBER: 60/063540  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063541  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063544  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 396  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-174-575A-396

Query Match 56.1%; Score 37; DB 6; Length 73;  
Best Local Similarity 72.7%; Pred. No. 1.5;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 3 SLSFXFTKFDL 13  
|||||  
Db 47 SLSFYFLKFQL 57

RESULT 14  
US-10-187-755-396  
; Sequence 396, Application US/10187755  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C1  
; CURRENT APPLICATION NUMBER: US/10/187,755  
; CURRENT FILING DATE: 2002-07-01  
; PRIOR APPLICATION NUMBER: US/10/052,586  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059266  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063120  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063121  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063486  
; PRIOR FILING DATE: 1997-10-21  
; PRIOR APPLICATION NUMBER: 60/063540  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063541  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063120  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063121  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063486  
; PRIOR FILING DATE: 1997-10-21  
; PRIOR APPLICATION NUMBER: 60/063540  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063541  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063544  
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 396  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-187-755-396

Query Match 56.1%; Score 37; DB 6; Length 73;  
Best Local Similarity 72.7%; Pred. No. 1.5;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDL 13  
|||||  
Db 47 SLSFYFLKFQL 57

RESULT 15  
US-10-187-749-396  
; Sequence 396, Application US/10187749  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C1  
; CURRENT APPLICATION NUMBER: US/10/187,749  
; CURRENT FILING DATE: 2002-07-01  
; PRIOR APPLICATION NUMBER: US/10/052,586  
; PRIOR FILING DATE: 2002-01-15  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059266  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063120  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063121  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/063486  
; PRIOR FILING DATE: 1997-10-21  
; PRIOR APPLICATION NUMBER: 60/063540  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063541  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/063544  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 396  
; LENGTH: 73  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-187-749-396

Query Match 56.1%; Score 37; DB 6; Length 73;  
Best Local Similarity 72.7%; Pred. No. 1.5;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 3 SLSFXFTKFDL 13  
|||||  
Db 47 SLSFYFLKFQL 57

Wed Feb 26 14:44:10 2003

us-09-476-485a-31.open.rapn

Page 6

Search completed: February 26, 2003, 14:55:22  
Job time : 60 secs

---

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:48:41 ; Search time 14 Seconds  
(without alignments)  
29.423 Million cell updates/sec

Title: US-09-476-485A-31  
Perfect score: 66  
Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTCUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description         |
|------------|-------|-------------|--------|-------|---------------------|
| 1          | 65    | 98.5        | 14     | 3     | US-08-825-369A-2    |
| 2          | 55    | 83.3        | 12     | 3     | US-08-825-369A-1    |
| 3          | 55    | 83.3        | 105    | 4     | US-08-881-189B-13   |
| 4          | 55    | 83.3        | 270    | 4     | US-08-881-189B-2    |
| 5          | 55    | 83.3        | 286    | 4     | US-08-881-189B-23   |
| 6          | 35    | 53.0        | 191    | 3     | US-08-655-352-5     |
| 7          | 35    | 53.0        | 191    | 3     | US-08-655-352-6     |
| 8          | 35    | 53.0        | 191    | 4     | US-09-258-016-5     |
| 9          | 35    | 53.0        | 191    | 4     | US-09-258-016-6     |
| 10         | 35    | 53.0        | 191    | 4     | US-09-257-825B-5    |
| 11         | 35    | 53.0        | 191    | 4     | US-09-257-825B-6    |
| 12         | 35    | 53.0        | 625    | 4     | US-09-134-001C-4504 |
| 13         | 34    | 51.5        | 187    | 1     | US-07-956-700B-17   |
| 14         | 34    | 51.5        | 187    | 1     | US-08-476-537-17    |
| 15         | 34    | 51.5        | 187    | 1     | US-08-485-607-17    |
| 16         | 34    | 51.5        | 187    | 2     | US-08-475-879-17    |
| 17         | 34    | 51.5        | 187    | 4     | US-09-433-043B-17   |
| 18         | 34    | 51.5        | 191    | 3     | US-08-655-352-7     |
| 19         | 34    | 51.5        | 191    | 4     | US-09-258-016-7     |
| 20         | 34    | 51.5        | 191    | 4     | US-09-257-825B-7    |
| 21         | 34    | 51.5        | 285    | 3     | US-09-141-821-2     |
| 22         | 34    | 51.5        | 285    | 3     | US-09-141-821-2     |
| 23         | 34    | 51.5        | 285    | 3     | US-09-141-821-4     |
| 24         | 34    | 51.5        | 285    | 3     | US-09-141-821-5     |
| 25         | 34    | 51.5        | 384    | 2     | US-08-833-226-2     |
| 26         | 34    | 51.5        | 449    | 2     | US-08-839-008-2     |
| 27         | 34    | 51.5        | 449    | 2     | US-08-839-008-9     |

28 34 51.5 491 1 US-07-956-700B-109 Sequence 109, App  
29 34 51.5 491 1 US-08-476-537-109 Sequence 109, App  
30 34 51.5 491 1 US-08-485-607-109 Sequence 109, App  
31 34 51.5 491 2 US-08-475-879-109 Sequence 109, App  
32 34 51.5 491 4 US-09-433-043B-109 Sequence 109, App  
33 34 51.5 491 4 US-09-433-043B-117 Sequence 117, App  
34 34 51.5 559 1 US-08-424-788-3 Sequence 3, Appli  
35 34 51.5 575 1 US-08-424-788-2 Sequence 2, Appli  
36 34 51.5 575 1 US-08-110-683-4 Sequence 4, Appli  
37 34 51.5 575 2 US-08-477-166-4 Sequence 4, Appli  
38 34 51.5 575 2 US-08-472-097-4 Sequence 4, Appli  
39 34 51.5 575 4 US-09-439-672-4 Sequence 4, Appli  
40 34 51.5 575 5 PCT-US93-11638-4 Sequence 4, Appli  
41 34 51.5 2172 1 US-08-611-107-31 Sequence 31, Appl  
42 34 51.5 2257 1 US-08-611-107-10 Sequence 10, Appl  
43 34 51.5 2257 2 US-08-422-560A-10 Sequence 10, Appl  
44 34 51.5 2257 4 US-08-468-793-10 Sequence 10, Appl  
45 33 50.0 169 4 US-08-720-625-4 Sequence 4, Appli

## ALIGNMENTS

RESULT 1  
US-08-825-369A-2  
; Sequence 2, Application US/08825369A  
; Patent No. 6084060  
; GENERAL INFORMATION:  
; APPLICANT: Moore  
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR PRESERVING  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron, LLP  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11753  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/825.369A  
; FILING DATE: March 28, 1997  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: O'Dea, Sean W.  
; REGISTRATION NUMBER: 37690  
; REFERENCE/DOCKET NUMBER: 381-21 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-825-369A-2

Query Match 98.5%; Score 65; DB 3; Length 14;  
Best Local Similarity 92.9%; Pred. No. 1.9e-06;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AQSLSFXFTKFDLD 14  
|||||

Db 1 AQSLSFSFTKFDLD 14  
|||||

RESULT 2

US-08-825-369A-1  
; Sequence 1, Application US/08825369A  
; Patent No. 6084060  
; GENERAL INFORMATION:  
; APPLICANT: Moore  
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR PRESERVING  
; TITLE OF INVENTION: PROGENITOR CELLS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron, LLP  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11753  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/825,369A  
; FILING DATE: March 28, 1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: O'Dea, Sean W.  
; REGISTRATION NUMBER: 37690  
; REFERENCE/DOCKET NUMBER: 381-21 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-825-369A-1

Query Match 83.3%; Score 55; DB 3; Length 12;  
Best Local Similarity 91.7%; Pred. No. 0.00013;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12  
| | | | | | | | | | | | | |  
DB 1 AQSLSFSFTKFD 12

RESULT 3  
US-08-881-189B-13  
; Sequence 13, Application US/08881189B  
; Patent No. 6310195  
; GENERAL INFORMATION:  
; APPLICANT: Colucci et al.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED  
; TITLE OF INVENTION: PROGENITOR CELL PRESERVATION FACTOR  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron, LLP  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11753  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/881,189B  
; FILING DATE: June 24, 1997  
; CLASSIFICATION: 424

Query Match 83.3%; Score 55; DB 4; Length 270;  
Best Local Similarity 91.7%; Pred. No. 0.0042;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: 381-44 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 105 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-881-189B-13

Query Match 83.3%; Score 55; DB 4; Length 105;  
Best Local Similarity 91.7%; Pred. No. 0.0015;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12  
| | | | | | | | | | | | | |  
DB 1 AQSLSFSFTKFD 12

RESULT 4  
US-08-881-189B-2  
; Sequence 2, Application US/08881189B  
; Patent No. 6310195  
; GENERAL INFORMATION:  
; APPLICANT: Colucci et al.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED  
; TITLE OF INVENTION: PROGENITOR CELL PRESERVATION FACTOR  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron, LLP  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11753  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/881,189B  
; FILING DATE: June 24, 1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: 381-44 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 270 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-881-189B-2

Query Match 83.3%; Score 55; DB 4; Length 270;  
Best Local Similarity 91.7%; Pred. No. 0.0042;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12  
| | | | | | | | | |  
Db 1 AQSLSFSFTKFD 12

RESULT 5  
US-08-881-189B-23  
; Sequence 23, Application US/08881189B  
; Patent No. 6310195  
; GENERAL INFORMATION:  
; APPLICANT: Colucci et al.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED  
; TITLE OF INVENTION: PROGENITOR CELL PRESERVATION FACTOR  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron, LLP  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11753  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/881,189B  
; FILING DATE: June 24, 1997  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Felt, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: 381-44 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 286 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-881-189B-23

Query Match 83.3%; Score 55; DB 4; Length 286;  
Best Local Similarity 91.7%; Pred. No. 0.0044;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12  
| | | | | | | | | |  
Db 23 AQSLSFSFTKFD 34

RESULT 6  
US-08-655-352-5  
; Sequence 5, Application US/08655352  
; Patent No. 6077991  
; GENERAL INFORMATION:  
; APPLICANT: Bachettira W. Poovaiah, Zhihua Liu,  
; APPLICANT: Shameekumar Patil, Daisuke Takezawa  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: PRODUCTION OF MALE-STERILE PLANTS  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klarquist Sparkman Campbell Leigh &  
; ADDRESSEE: Whinston, LLP  
; STREET: One World Trade Center  
; STREET: 121 S.W. Salmon Street  
; STREET: Suite 1600

; CITY: Portland  
; STATE: Oregon  
; COUNTRY: United States of America  
; ZIP: 97204  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Disk, 3-1/2 inch  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: MS DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/655,352  
; FILING DATE:  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/323,449  
; FILING DATE: October 14, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dow, Alan. E.  
; REGISTRATION NUMBER: 35,123  
; REFERENCE/DOCKET NUMBER: 4630-45000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (503) 226-7391  
; TELEFAX: (503) 228-9446  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 191 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; DESCRIPTION: Region of rat neural visinin-like protein  
; DESCRIPTION: (Gen2:Ratnvp1) with homology to lily  
; DESCRIPTION: CCaMK  
US-08-655-352-5

Query Match 53.0%; Score 35; DB 3; Length 191;  
Best Local Similarity 46.2%; Pred. No. 19;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 QSLSFXFTKFDLD 14  
| | | | | | | | | |  
Db 99 QKLNWAFNMYDLD 111

RESULT 7  
US-08-655-352-6  
; Sequence 6, Application US/08655352  
; Patent No. 6077991  
; GENERAL INFORMATION:  
; APPLICANT: Bachettira W. Poovaiah, Zhihua Liu,  
; APPLICANT: Shameekumar Patil, Daisuke Takezawa  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: PRODUCTION OF MALE-STERILE PLANTS  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klarquist Sparkman Campbell Leigh &  
; ADDRESSEE: Whinston, LLP  
; STREET: One World Trade Center  
; STREET: 121 S.W. Salmon Street  
; STREET: Suite 1600  
; CITY: Portland  
; STATE: Oregon  
; COUNTRY: United States of America  
; ZIP: 97204  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Disk, 3-1/2 inch  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: MS DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/655,352  
; FILING DATE:  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:



APPLICANT: Poovaiah, Bachettira W.  
APPLICANT: Patil, Shameekumar  
APPLICANT: Takezawa, Daisuke  
TITLE OF INVENTION: Compositions and Methods for Production of Male-Sterile Plants  
FILE REFERENCE: 4630-51993  
CURRENT APPLICATION NUMBER: US/09/257,825B  
CURRENT FILING DATE: 1999-02-25  
PRIOR APPLICATION NUMBER: US 08/655,352  
PRIOR FILING DATE: 1996-05-23  
PRIOR APPLICATION NUMBER: US 60/014,743  
PRIOR FILING DATE: 1996-03-28  
NUMBER OF SEQ ID NOS: 28  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 5  
LENGTH: 191  
TYPE: PRT  
ORGANISM: Rat  
US-09-257-825B-5

Query Match 53.0%; Score 35; DB 4; Length 191;  
Best Local Similarity 46.2%; Pred. No. 19;  
Matches 6; Conservative 3; Mismatches 0; Indels 4; Gaps 0;

QY 2 QSLSFXTKFDLD 14  
| : : | : |||  
Db 99 QKLNWAFNMYDLD 111

RESULT 11  
US-09-257-825B-6  
; Sequence 6, Application US/09257825B  
; Patent No. 6403352  
; GENERAL INFORMATION:  
; APPLICANT: Poovaiah, Bachettira W.  
; APPLICANT: Patil, Shameekumar  
; APPLICANT: Takezawa, Daisuke  
; TITLE OF INVENTION: Compositions and Methods for Production of Male-Sterile Plants  
; FILE REFERENCE: 4630-51993  
; CURRENT APPLICATION NUMBER: US/09/257,825B  
; CURRENT FILING DATE: 1999-02-25  
; PRIOR APPLICATION NUMBER: US 08/655,352  
; PRIOR FILING DATE: 1996-05-23  
; PRIOR APPLICATION NUMBER: US 60/014,743  
; PRIOR FILING DATE: 1996-03-28  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 191  
; TYPE: PRT  
; ORGANISM: Chicken  
US-09-257-825B-6

Query Match 53.0%; Score 35; DB 4; Length 191;  
Best Local Similarity 46.2%; Pred. No. 19;  
Matches 6; Conservative 3; Mismatches 0; Indels 4; Gaps 0;

QY 2 QSLSFXTKFDLD 14  
| : : | : |||  
Db 99 QKLNWAFNMYDLD 111

RESULT 12  
US-09-134-001C-4504  
; Sequence 4504, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 4504  
; LENGTH: 625  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4504

Query Match 53.0%; Score 35; DB 4; Length 625;  
Best Local Similarity 60.0%; Pred. No. 69;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 LSFXTKFDL 13  
| : : | : |||  
Db 265 ISYHFNKFDL 274

RESULT 13  
US-07-956-700B-17  
; Sequence 17, Application US/07956700B  
; Patent No. 5539092  
; GENERAL INFORMATION:  
; APPLICANT: Robert Haselkorn and Piotr Gornicki  
; TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA  
; NUMBER OF SEQUENCES: 116  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60610

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII-DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/956,700B  
FILING DATE: 19921002  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Thomas E. No. 5539092thrup  
REGISTRATION NUMBER: 33,268  
REFERENCE/DOCKET NUMBER: ARCD:058  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 1-312-744-0090  
TELEFAX: 1-312-755-4489  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 187 amino acids  
TYPE: Amino acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Peptide  
US-07-956-700B-17

Query Match 51.5%; Score 34; DB 1; Length 187;  
Best Local Similarity 53.8%; Pred. No. 28;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 QSLSFXTKFDLD 14  
| : : | : |||  
Db 149 KEISAVATKFDLD 161

RESULT 14  
US-08-476-537-17  
; Sequence 17, Application US/08476537  
; Patent No. 5756290

GENERAL INFORMATION:  
APPLICANT: Robert Haselkorn and Piotr Gornicki  
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA  
TITLE OF INVENTION: Carboxylase  
NUMBER OF SEQUENCES: 116  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: 321 No. 5756290th Clark Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60610  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII-DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,537  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/956,700  
FILING DATE: 10/21/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Thomas E. No. 5756290thrup  
REGISTRATION NUMBER: 33,268  
REFERENCE/DOCKET NUMBER: ARCD:058  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 1-312-744-0090  
TELEFAX: 1-312-755-4489  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 187 amino acids  
TYPE: Amino acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Peptide  
US-08-476-537-17

Query Match 51.5%; Score 34; DB 1; Length 187;  
Best Local Similarity 53.8%; Pred. No. 28;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 QSLSFXTKFDLD 14  
Db 149 KEISAVATKFDLD 161

RESULT 15  
US-08-485-607-17  
Sequence 17, Application US/08485607  
Patent No. 5792627  
GENERAL INFORMATION:  
APPLICANT: Robert Haselkorn and Piotr Gornicki  
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA  
TITLE OF INVENTION: Carboxylase  
NUMBER OF SEQUENCES: 116  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: 321 No. 5792627th Clark Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60610  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII-DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,607  
FILING DATE: 07-JUN-1995

CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/956,700  
FILING DATE: 10/21/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Thomas E. No. 5792627thrup  
REGISTRATION NUMBER: 33,268  
REFERENCE/DOCKET NUMBER: ARCD:058  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 1-312-744-0090  
TELEFAX: 1-312-755-4489  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 187 amino acids  
TYPE: Amino acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Peptide  
US-08-485-607-17

Query Match 51.5%; Score 34; DB 1; Length 187;  
Best Local Similarity 53.8%; Pred. No. 28;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 QSLSFXTKFDLD 14  
Db 149 KEISAVATKFDLD 161

Search completed: February 26, 2003, 14:51:13  
Job time : 14 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:42:12 ; Search time 14 Seconds  
(without alignments)  
54.934 Million cell updates/sec

Title: US-09-476-485A-24  
Perfect score: 36  
Sequence: 1 TNNVLOXT 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 604

Minimum DB seq length: 0  
Maximum DB seq length: 8

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description         |
|------------|-------|-------------|--------|----------|---------------------|
| 1          | 12    | 33.3        | 5      | 2 B31836 | 20K protein - Rick  |
| 2          | 12    | 33.3        | 7      | 2 S78024 | ribosomal protein   |
| 3          | 12    | 33.3        | 8      | 2 S70727 | ip9F protein - Shi  |
| 4          | 12    | 33.3        | 8      | 2 C61512 | variant surface gl  |
| 5          | 12    | 33.3        | 8      | 2 A14683 | aspartate transami  |
| 6          | 11    | 30.6        | 5      | 2 PT0295 | Ig heavy chain CRD  |
| 7          | 11    | 30.6        | 6      | 2 S14159 | parasporal crystal  |
| 8          | 11    | 30.6        | 6      | 2 B31263 | dihydrofolate redu  |
| 9          | 11    | 30.6        | 7      | 2 S45311 | microcin C7 - Esch  |
| 10         | 11    | 30.6        | 8      | 2 PL0184 | capsid protein VP-  |
| 11         | 11    | 30.6        | 8      | 2 S68802 | nitrate reductase   |
| 12         | 10    | 27.8        | 4      | 2 A37832 | phenol 2-monooxyge  |
| 13         | 10    | 27.8        | 7      | 2 I40504 | hypothetical prote  |
| 14         | 10    | 27.8        | 7      | 2 A34026 | acetylcholinestera  |
| 15         | 10    | 27.8        | 8      | 2 S43971 | tumor-associated a  |
| 16         | 10    | 27.8        | 8      | 2 S43972 | tumor-associated a  |
| 17         | 10    | 27.8        | 8      | 2 A46306 | spasmogenic toxin   |
| 18         | 10    | 27.8        | 8      | 2 A47618 | beta-galactosidase  |
| 19         | 9     | 25.0        | 4      | 2 A48360 | gamma subunit of p  |
| 20         | 9     | 25.0        | 5      | 2 A37114 | hypoxanthine phosph |
| 21         | 9     | 25.0        | 5      | 2 PT0625 | T-cell receptor be  |
| 22         | 9     | 25.0        | 6      | 2 JH0784 | neuropeptide TE-6   |
| 23         | 9     | 25.0        | 6      | 2 PT0560 | T-cell receptor be  |
| 24         | 9     | 25.0        | 7      | 1 XE1DGD | galactose oxidase   |
| 25         | 9     | 25.0        | 7      | 2 S19630 | ribosomal protein   |
| 26         | 9     | 25.0        | 7      | 2 B44787 | callifMRamide 11    |
| 27         | 9     | 25.0        | 7      | 2 S68004 | hucolin, 75K chain  |
| 28         | 9     | 25.0        | 7      | 2 PT0246 | Ig heavy chain CRD  |
| 29         | 9     | 25.0        | 7      | 2 I48086 | DNA topoisomerase   |

serine/threonine-s  
polyphosphate-gluc  
hypertrehalosemic  
adipokinetic hormo  
neuropeptide led-C  
adipokinetic hormo  
adipokinetic hormo  
hypothetical prote  
protein QA300039 -  
callifMRamide 8 -  
cytochrome-c oxida  
major postsynaptic  
D-mannanate hydrol  
branched-chain-ami  
zinc-binding prote  
Ig heavy chain CRD

## ALIGNMENTS

## RESULT 1

B31836  
20K protein - Rickettsia rickettsii (fragment)  
C:Species: Rickettsia rickettsii  
C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 28-May-1999  
C:Accession: B31836  
R:Anderson, B.E.; Baumstark, B.R.; Bellini, W.J.  
J. Bacteriol. 170, 4493-4500, 1988  
A:Title: Expression of the gene encoding the 17-kilodalton antigen from Rickettsia rickettsii  
A:Reference number: A91885; MUID:89008059; PMID:3139629  
A:Accession: B31836  
A:Molecule type: DNA  
A:Residues: 1-5 <AND>  
A:Cross-references: GB:J03371; NID:gl52455; PIDN:AAI5030.1; PID:g4262874

Query Match 33.3%; Score 12; DB 2; Length 5;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNN 3  
||:  
Db 3 TNS 5

## RESULT 2

S78024  
ribosomal protein YmL21, mitochondrial - yeast (Saccharomyces cerevisiae) (fragment)  
C:Species: Saccharomyces cerevisiae  
C:Date: 09-Oct-1997 #sequence\_revision 24-Oct-1997 #text\_change 14-Nov-1997  
C:Accession: S78024  
R:Kitakawa, M.; Graack, H.R.; Grohmann, L.; Goldschmidt-Reisin, S.; Herfurth, E.; Wit  
Eur. J. Biochem. 245, 449-456, 1997  
A:Title: Identification and characterization of the genes for mitochondrial ribosomal  
A:Reference number: S78018; MUID:97296414; PMID:9151978  
A:Accession: S78024  
A:Molecule type: protein  
A:Residues: 1-7 <KIT>  
C:Genetics:  
A:Genome: nuclear  
C:Keywords: mitochondrion; protein biosynthesis; ribosome

Query Match 33.3%; Score 12; DB 2; Length 7;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLIQ 6  
|:|:  
Db 4 NLK 7

## RESULT 3

S70727

ipgF protein - Shigella flexneri (fragment)  
 C:Species: Shigella flexneri  
 C:Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 08-Oct-1999  
 C:Accession: S70727  
 R:Aliaoui, A.; Sansonetti, P.J.; Menard, R.; Barzu, S.; Mounier, J.; Phalipon, A.; Parsc  
 Mol. Microbiol. 17, 461-470, 1995  
 A:Title: MxiG, a membrane protein required for secretion of Shigella spp. Ipa invasins:  
 A:Reference number: S70727; MUID:96100445; PMID:8559065  
 A:Accession: S70727  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-8 <ALL>  
 A:Cross-references: EMBL:Z48957; NID:929880; PIDN:CAA88821.1; PID:g929881  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1995  
 C:Genetics:  
 A:Gene: ipgF

Query Match 33.3%; Score 12; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NN 3  
 ||  
 Db 4 NN 5

RESULT 4  
 C61512  
 Variant surface glycoprotein MITat 1.4 - Trypanosoma brucei (fragment)  
 C:Species: Trypanosoma brucei  
 C:Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 07-May-1999  
 C:Accession: C61512  
 R:Holder, A.A.; Cross, G.A.M.  
 Mol. Biochem. Parasitol. 2, 135-150, 1981  
 A:Title: Glycopeptides from variant surface glycoproteins of Trypanosoma brucei. C-termi  
 A:Reference number: A61512; MUID:81172836; PMID:6163983  
 A:Accession: C61512  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-8 <HOL>  
 C:Keywords: glycoprotein

Query Match 33.3%; Score 12; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NN 3  
 ||  
 Db 3 NN 4

RESULT 5  
 A14683  
 aspartate transaminase (EC 2.6.1.1), mitochondrial - chicken (tentative sequence) (fragm  
 N:Alternate names: aspartate aminotransferase, mitochondrial  
 C:Species: Gallus gallus (chicken)  
 C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 31-Mar-2000  
 C:Accession: A14683  
 R:Wilson, K.J.; Hunziker, P.; Hughes, G.J.  
 FEBS Lett. 108, 98-102, 1979  
 A:Title: Microsequence analysis: IV. Automatic liquid-phase sequencing using DABITC.  
 A:Reference number: A14683; MUID:80092116; PMID:520566  
 A:Accession: A14683  
 A:Molecule type: protein  
 A:Residues: 1-8 <WIL>  
 C:Keywords: aminotransferase; mitochondrion

Query Match 33.3%; Score 12; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NN 3  
 ||

Db 3 NN 4  
 RESULT 6  
 PT0295  
 Ig heavy chain CRD3 region (clone 5-91) - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
 C:Accession: PT0295  
 R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
 J. Exp. Med. 173, 395-407, 1991  
 A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity an  
 A:Reference number: PT0222; MUID:91108337; PMID:1899102  
 A:Accession: PT0295  
 A:Molecule type: DNA  
 A:Residues: 1-5 <YAM>  
 A:Experimental source: B lymphocyte  
 C:Keywords: heterotetramer; immunoglobulin

Query Match 30.6%; Score 11; DB 2; Length 5;  
 Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNN 3  
 | |  
 Db 3 TRN 5

RESULT 7  
 S14159  
 Parasporal crystal protein CryIC - Bacillus thuringiensis (fragment)  
 N:Alternate names: delta-endotoxin  
 C:Species: Bacillus thuringiensis  
 C:Date: 19-Mar-1997 #sequence\_revision 18-Jul-1997 #text\_change 18-Jul-1997  
 C:Accession: S14159  
 R:Convents, D.; Cherlet, M.; van Damme, J.; Lasters, I.; Lauwereys, M.  
 Eur. J. Biochem. 195, 631-635, 1991  
 A:Title: Two structural domains as a general fold of the toxic fragment of the Bacill  
 A:Reference number: S14087; MUID:91153300; PMID:1847865  
 A:Accession: S14159  
 A:Molecule type: protein  
 A:Residues: 1-6 <CON>

Query Match 30.6%; Score 11; DB 2; Length 6;  
 Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNN 3  
 | |  
 Db 3 TGN 5

RESULT 8  
 B31263  
 dihydrofolate reductase (EC 1.5.1.3) / thymidylate synthase (EC 2.1.1.45) - Plasmodiu  
 C:Species: Plasmodium falciparum  
 C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 20-Mar-1996  
 C:Accession: B31263  
 R:Peterson, D.S.; Walliker, D.; Wellem, T.E.  
 Proc. Natl. Acad. Sci. U.S.A. 85, 9114-9118, 1988  
 A:Title: Evidence that a point mutation in dihydrofolate reductase-thymidylate syntha  
 A:Reference number: A94217; MUID:89057886; PMID:2904149  
 A:Accession: B31263  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-6 <PET>  
 C:Keywords: methyltransferase; NADP; oxidoreductase

Query Match 30.6%; Score 11; DB 2; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TN 2

Db 1 TN 2  
||

## RESULT 9

S45311 microcin C7 - Escherichia coli plasmid pmccc7

C:Species: Escherichia coli

C:Date: 10-Dec-1994 #sequence\_revision 24-May-1996 #text\_change 17-Mar-1999

C:Accession: S45311

R:Gonzalez-Pastor, J.E.; San Millan, J.L.; Moreno, F.

Nature 369, 281, 1994

A:Title: The smallest known gene.

A:Reference number: S45311; MUID:94239518; PMID:8183363

A:Accession: S45311

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-7 <GON>

C:Genetics:

A:Gene: mcccA

A:Genome: plasmid pmccc7

C:Keywords: antibacterial

F:1/Modified site: N-formylmethionine #status predicted

F:7/Modified site: asparagine derivative (Asn) #status experimental

Query Match

Best Local Similarity 30.6%; Score 11; DB 2; Length 7;

Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNN 3

||

Db 3 TGN 5

## RESULT 10

PL0184

capsid protein VP-1 - murine poliovirus (fragment)

C:Species: murine poliovirus, Theiler's encephalomyelitis virus

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 20-Feb-1995

C:Accession: PL0184

R:Zurbriggen, A.; Hoque, J.M.; Fujinami, R.S.

J. Exp. Med. 170, 2037-2049, 1989

A:Title: Alteration of amino acid 101 within capsid protein VP-1 changes the pathogenicity

A:Reference number: PL0184; MUID:90063468; PMID:2479706

A:Accession: PL0184

A:Molecule type: genomic RNA

A:Residues: 1-8 <ZUR>

C:Keywords: capsid protein

Query Match

Best Local Similarity 30.6%; Score 11; DB 2; Length 8;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TN 2

||

Db 5 TN 6

## RESULT 11

S68802

nitrate reductase (NADH) inhibitor - spinach (fragment)

C:Species: Spinacia oleracea (spinach)

C:Date: 04-Dec-1997 #sequence\_revision 12-Dec-1997 #text\_change 30-Jan-1998

C:Accession: S68802

R:Bachmann, M.; Huber, J.L.; Liao, P.C.; Gage, D.A.; Huber, S.C.

FEBS Lett. 387, 127-131, 1996

A:Title: The inhibitor protein of phosphorylated nitrate reductase from spinach (Spinacia

A:Reference number: S68802; MUID:96244508; PMID:8674533

A:Accession: S68802

A:Molecule type: protein

A:Residues: 1-8 <BAC>

A:Experimental source: leaves; strain cv. Bloomsdale

Query Match 30.6%; Score 11; DB 2; Length 8;  
Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVL 5

||

Db 1 NLL 3

## RESULT 12

A37832

phenol 2-monooxygenase (EC 1.14.13.7) chain P5 - Pseudomonas sp. (strain CF600) (frag

C:Species: Pseudomonas sp.

C:Date: 14-Jun-1991 #sequence\_revision 14-Jun-1991 #text\_change 23-Jun-1993

C:Accession: A37832

R:Powlowski, J.; Shingler, V.

J. Bacteriol. 172, 6834-6840, 1990

A:Title: In vitro analysis of polypeptide requirements of multicomponent phenol hydro

A:Reference number: A37832; MUID:91072231; PMID:2254259

A:Accession: A37832

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-4 <POW>

C:Keywords: oxidoreductase

Query Match

Best Local Similarity 27.8%; Score 10; DB 2; Length 4;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NV 4

||

Db 3 NV 4

## RESULT 13

I40504

hypothetical protein 2 (7 aa) - Bacillus stearothermophilus

C:Species: Bacillus stearothermophilus

C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 15-Oct-1999

C:Accession: I40504

R:Waye, M.M.; Winter, G.

Eur. J. Biochem. 158, 505-510, 1986

A:Title: A transcription terminator in the 5' non-coding region of the tyrosyl tRNA s

A:Reference number: I40503; MUID:86274732; PMID:3525162

A:Accession: I40504

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-7 <RES>

A:Cross-references: EMBL:X04193; MUID:940233; PIDN:CAA27782.1; PID:9580943

Query Match

Best Local Similarity 27.8%; Score 10; DB 2; Length 7;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NV 4

||

Db 3 NV 4

## RESULT 14

A34026

acetylcholinesterase (EC 3.1.1.7) 5.6S form - Pacific electric ray (fragment)

C:Species: Torpedo californica (Pacific electric ray)

C:Date: 16-Mar-1990 #sequence\_revision 16-Mar-1990 #text\_change 08-Nov-1996

C:Accession: A34026

R:Gibney, G.; MacPhee-Quigley, K.; Thompson, B.; Vedvick, T.; Low, M.G.; Taylor, S.S.

J. Biol. Chem. 263, 1140-1145, 1988

A:Title: Divergence in primary structure between the molecular forms of acetylcholine

A:Reference number: A34026; MUID:88087329; PMID:3355334

A:Accession: A34026

A:Molecule type: protein

A:Residues: 1-7 <GIB>

C:Keywords: alternative splicing; carboxylic ester hydrolase

Query Match 27.8%; Score 10; DB 2; Length 7;  
Best Local Similarity 40.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 VLQXT 8  
: | |  
Db 1 LLNAT 5

## RESULT 15

S43971  
tumor-associated antigen MUR1 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 20-Oct-1994 #sequence\_revision 17-Nov-1995 #text\_change 18-Aug-2000  
C:Accession: S43971  
R:Mandelblat, O.; Berke, G.; Fridkin, M.; Feldman, M.; Eisenstein, M.; Eisenbach, L.  
Nature 369, 67-71, 1994  
A:Title: CTL induction by a tumour-associated antigen octapeptide derived from a murine  
A:Reference number: S43971; MUID:94217811; PMID:8164742  
A:Accession: S43971  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-8 <MAN>  
C:Superfamily: unassigned animal peptides

Query Match 27.8%; Score 10; DB 2; Length 8;  
Best Local Similarity 50.0%; Pred. No. 2.8e+05;  
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NVLQ 6  
: | |  
Db 4 NTAQ 7

Search completed: February 26, 2003, 15:45:43  
Job time : 15 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:39:16 ; Search time 11 seconds  
(without alignments)  
30.165 Million cell updates/sec

Title: US-09-476-485A-24  
Perfect score: 36  
Sequence: 1 TNNVLOXT 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 150

Minimum DB seq length: 0

Maximum DB seq length: 8

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID           | Description        |
|------------|-------|-------------|--------|--------------|--------------------|
| 1          | 12    | 33.3        | 5      | 1 TPIS_CANFA | P54714 canis famil |
| 2          | 12    | 33.3        | 8      | 1 FUSS_FUSSO | P81010 fusarium so |
| 3          | 11    | 30.6        | 6      | 1 TMOF_SARBU | P41495 sarcophaga  |
| 4          | 11    | 30.6        | 6      | 1 VP19_HSVIK | P23210 herpes simp |
| 5          | 11    | 30.6        | 8      | 1 FAR4_HOMAM | P41487 homarus ame |
| 6          | 11    | 30.6        | 8      | 1 UPAL_HUMAN | P30087 homo sapien |
| 7          | 10    | 27.8        | 7      | 1 FAR1_PROCL | P38499 procambarus |
| 8          | 10    | 27.8        | 7      | 1 FAR2_PROCL | P38498 procambarus |
| 9          | 10    | 27.8        | 8      | 1 FAR1_PANRE | P41872 panagrellus |
| 10         | 10    | 27.8        | 8      | 1 FAR3_HOMAM | P41866 homarus ame |
| 11         | 9     | 25.0        | 7      | 1 FAR6_CALVO | P41866 calliphora  |
| 12         | 9     | 25.0        | 7      | 1 IGAO_DACDE | P06294 dactylum d  |
| 13         | 9     | 25.0        | 8      | 1 AKH_TABAT  | P14595 tabanus atr |
| 14         | 9     | 25.0        | 8      | 1 B44K_PORGI | P81886 porphyromon |
| 15         | 9     | 25.0        | 8      | 1 COXC_RAT   | P80430 rattus norv |
| 16         | 9     | 25.0        | 8      | 1 FAR6_CALVO | P41863 calliphora  |
| 17         | 9     | 25.0        | 8      | 1 HTF2_PERAM | P04549 periplaneta |
| 18         | 8     | 22.2        | 6      | 1 ACPH_RABIT | P25154 oryctolagus |
| 19         | 8     | 22.2        | 7      | 1 FAR1_ASCSU | P31889 ascaris suu |
| 20         | 8     | 22.2        | 7      | 1 FAR4_PANRE | P41875 panagrellus |
| 21         | 8     | 22.2        | 7      | 1 MNPI_LEPDE | P42984 leptinotars |
| 22         | 8     | 22.2        | 8      | 1 CADL_ENTFA | P13268 enterococcu |
| 23         | 8     | 22.2        | 8      | 1 CLP_THICU  | P80488 thiobacillu |
| 24         | 8     | 22.2        | 8      | 1 COW2_CONPU | P58785 conus purpu |
| 25         | 7     | 19.4        | 5      | 1 BIOA_CITFR | P13071 citrobacter |
| 26         | 7     | 19.4        | 5      | 1 TRM3_ECOLI | P13973 escherichia |
| 27         | 7     | 19.4        | 6      | 1 AS2P_LACSN | P82655 lactobacill |
| 28         | 7     | 19.4        | 6      | 1 UN06_CLOPA | P81351 clostridium |
| 29         | 7     | 19.4        | 7      | 1 CHOX_ALCSP | P16101 alcaligenes |
| 30         | 7     | 19.4        | 7      | 1 CIA_ENTFA  | P11932 enterococcu |
| 31         | 7     | 19.4        | 7      | 1 GFRE_MOUSE | P99025 mus musculu |
| 32         | 7     | 19.4        | 7      | 1 LANC_CARUI | P36960 carnobacter |
| 33         | 7     | 19.4        | 7      | 1 UN06_PINPS | P81675 pinus pinas |

34 7 19.4 8 1 LCK1\_LEUMA  
35 7 19.4 8 1 LCK3\_LEUMA  
36 7 19.4 8 1 PLP\_BRANA  
37 7 19.4 8 1 UPAA\_HUMAN  
38 6 16.7 4 1 FLRN\_ANTEL  
39 6 16.7 6 1 FARP\_MONEX  
40 6 16.7 7 1 ALL5\_CARMA  
41 6 16.7 7 1 CARP\_MYTEZ  
42 6 16.7 7 1 UC24\_MAIZE  
43 6 16.7 8 1 ACI\_THUAL  
44 6 16.7 8 1 AKHG\_GRYBI  
45 6 16.7 8 1 AKH\_LIBAU

P21140 leucophaea  
P21142 leucophaea  
P81707 brassica na  
P30096 homo sapien  
P58707 anthopleura  
P41966 moniezia ex  
P81808 carcinus ma  
P10420 mytilus edu  
P80630 zea mays (m  
P18691 thunnus alb  
P14086 gryllus bim  
P25418 libellula a

## ALIGNMENTS

RESULT 1  
TPIS\_CANFA STANDARD; PRT; 5 AA.  
AC P54714;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Triosephosphate isomerase (EC 5.3.1.1) (TIM) (Fragment).  
GN tpil.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Heart;  
RX MEDLINE=98163340; PubMed=9504812;  
RA Dunn M.J., Corbett J.M., Wheeler C.H.;  
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of  
RT dog heart proteins.";  
RL Electrophoresis 18:2795-2802(1997).  
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate -> glycerone  
CC phosphate.  
CC -!- PATHWAY: PLAYS AN IMPORTANT ROLE IN SEVERAL METABOLIC PATHWAYS.  
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.  
DR HSC-2DPAGE; P54714; DOG.  
DR InterPro: IPR000652; Triophos\_ismrse.  
DR PROSITE: PS00171; TIM; PARTIAL.  
KW Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;  
KW Pentose shunt.  
FT NON\_TER 1 1  
FT NON\_TER 5 5  
SQ SEQUENCE 5 AA; 550 MW; 64444862C9A00000 CRC64;  
Query Match 33.3%; Score 12; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 NN 3  
Db 4 NN 5  
RESULT 2  
FUSS\_FUSSO STANDARD; PRT; 8 AA.  
AC P81010;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Allergen Fus s I3596\* (Fragment).  
OS Fusarium solani (subsp. pisi) (Nectria haematococca).  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Hypocreales; Nectriaceae; Nectria.  
OX NCBI\_TaxID=70791;

```

RN SEQUENCE.
RP STRAIN-IARI 3596; TISSUE=Myelium;
RA Verma J., Gangal S.V.;
RL Submitted (JUL-1997) to the SWISS-PROT data bank.
KW Allergen.
FT NON_TER
SQ SEQUENCE 8 AA; 898 MW; C372C441F5B69041 CRC64;

Query Match 33.3%; Score 12; DB 1; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.1e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4
DB 4 SHNV 7

RESULT 3
TMOF_SARBU STANDARD; PRT; 6 AA.
AC P41495.
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Trypsin-modulating oostatic factor (TMOF).
OS Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Oestroidea; Sarcophagidae; Sarcophaga.
OX NCBI_TaxID=7385;
RN [1]
RC TISSUE=Ovary;
RX MEDLINE=94211930; PubMed=8159807;
RA Bylemans D., Borovsky D., Hunt D.F., Shabanowitz J., Grauwels L.,
RA de Loof A.;
RT "Sequencing and characterization of trypsin modulating oostatic
RT factor (TMOF) from the ovaries of the grey fleshfly, Neobellieria
RT (Sarcophaga) bullata."
RL Regul. Pept. 50:61-72(1994).
CC -!- FUNCTION: HAS AN OOSTATIC ACTIVITY. INHIBITS TRYPSIN BIOSYNTHESIS
CC IN THE MIDGUT WHICH INDIRECTLY REDUCES THE VITELLOGENIN
CC CONCENTRATION IN THE HEMOLYMPH RESULTING IN INHIBITION OF OOCYTE
CC DEVELOPMENT.
CC -!- DEVELOPMENTAL STAGE: SYNTHESIZED AND RELEASED FROM FOLLICULAR
CC EPITHELIUM AFTER A BLOOD MEAL.
KW Hormone.
SQ SEQUENCE 6 AA; 695 MW; 61E72451B7642000 CRC64;

Query Match 30.6%; Score 11; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TN 2
DB 3 TN 4

RESULT 4
VP19_HSVIK STANDARD; PRT; 6 AA.
AC P23210;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Capsid assembly and DNA maturation protein (Virion protein UL38)
DE (Capsid protein VP19C) (Fragment).
GN UL38.
OS Herpes simplex virus (type 1 / strain KOS).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10306;

```

```

RN SEQUENCE FROM N.A.
RP MEDLINE=91101287; PubMed=1846198;
RA Flanagan W.M., Papavasiliou A.G., Rice M., Hecht L.B.,
RA Silverstein S., Wagner E.K.;
RT "Analysis of the herpes simplex virus type 1 promoter controlling the
RT expression of UL38, a true late gene involved in capsid assembly."
RL J. Virol. 65:769-786(1991).
CC -!- FUNCTION: COMPONENT OF THE BASAL LAYER IN WHICH THE CAPSIDS ARE
CC EMBEDDED. BINDS DNA.
CC -!- SIMILARITY: TO OTHER HERPESVIRUSES CAPSID PROTEIN VP19C.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M57646; AAA45830.1; -
KW Capsid assembly; Coat protein; DNA-binding.
FT NON_TER
SQ SEQUENCE 6 AA; 703 MW; 67376451A336F000 CRC64;

Query Match 30.6%; Score 11; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TN 2
DB 3 TN 4

RESULT 5
FAR4_HOMAM STANDARD; PRT; 8 AA.
ID FAR4_HOMAM
AC P41487;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FMRFamide-like neuropeptide 4 (FLI 4) (FL).
OS Homarus americanus (American lobster).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
OC Astacidea; Nephropoidea; Nephropidae; Homarus.
OX NCBI_TaxID=6706;
RN [1]
RC TISSUE=Pericardial organs;
RX MEDLINE=8816164; PubMed=3429714;
RA Trimmer B.A., Kobierski L.A., Kravitz E.A.;
RT "Purification and characterization of FMRFamide-like immunoreactive
RT substances from the lobster nervous system: isolation and sequence
RT analysis of two closely related peptides."
RL J. Comp. Neurol. 266:16-26(1987)
CC -!- FUNCTION: CAN ACT AS A MODULATOR OF EXOSKELETAL AND CARDIAC
CC NEUROMUSCULAR JUNCTIONS.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES
SQ SEQUENCE 8 AA; 1067 MW; DDD40729C4540451 CRC64;

Query Match 30.6%; Score 11; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TN 2
DB 1 TN 2

```

```

RESULT 6
ID UPAL_HUMAN STANDARD; PRT; 8 AA.
AC P30087;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Unknown protein from 2D-page of plasma (Spot 2) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RC TISSUE=Plasma;
RX MEDLINE=93092937; PubMed=1459097;
RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
RA Hochstrasser D.F.;
RT "Plasma protein map: an update by microsequencing.";
RL Electrophoresis 13:707-714(1992)
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 4.9, ITS MW IS: 65 Kda.
DR SWISS-2DPAGE; P30087; HUMAN.
FT NON_TER 1
FT UNSURE 8
FT NON_TER 8
SQ SEQUENCE 8 AA; 944 MW; C01772C455BB06DA CRC64;

Query Match 30.6%; Score 11; DB 1; Length 8;
Best Local Similarity 66.7%; Pred. No. 1.1e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 NN 4
DB 4 SN 6

RESULT 7
ID FAR1_PROCL STANDARD; PRT; 7 AA.
AC P38499;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Cardioexcitatory FMRFamide homolog NFL.
OS Procamburus clarkii (Red swamp crayfish).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
OC Astacidea; Astacoidea; Cambaridae; Procamburus.
OX NCBI_TaxID=6728;
RN [1]
RP SEQUENCE.
RC TISSUE=Pericardial organs;
RX MEDLINE=93248032; PubMed=8387183;
RA Mercier A.J., Orchard I., Tebrugge V., Skerrett M.;
RT "Isolation of two FMRFamide-related peptides from crayfish
RT pericardial organs.";
RL Peptides 14:137-143(1993)
CC -!- FUNCTION: INCREASES THE RATE AND AMPLITUDE OF SPONTANEOUS
CC CONTRACTIONS OF SEMI-ISOLATED HEARTS. INCREASES THE AMPLITUDE OF
CC EXCITATORY POSTSYNAPTIC POTENTIALS IN ABDOMINAL EXTENSOR MUSCLE.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 7
FT SEQUENCE 7 AA; 967 MW; 69D40729C4540AC0 CRC64;

Query Match 27.8%; Score 10; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.1e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 NVLQ 6
DB 3 NFLR 6

RESULT 8
ID FAR2_PROCL STANDARD; PRT; 7 AA.
AC P38498;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Cardioexcitatory FMRFamide homolog DF2
OS Procamburus clarkii (Red swamp crayfish).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
OC Astacidea; Astacoidea; Cambaridae; Procamburus.
OX NCBI_TaxID=6728;
RN [1]
RP SEQUENCE.
RC TISSUE=Pericardial organs;
RX MEDLINE=93248032; PubMed=8387183;
RA Mercier A.J., Orchard I., Tebrugge V., Skerrett M.;
RT "Isolation of two FMRFamide-related peptides from crayfish
RT pericardial organs.";
RL Peptides 14:137-143(1993)
CC -!- FUNCTION: INCREASES THE RATE AND AMPLITUDE OF SPONTANEOUS
CC CONTRACTIONS OF SEMI-ISOLATED HEARTS. INCREASES THE AMPLITUDE OF
CC EXCITATORY POSTSYNAPTIC POTENTIALS IN ABDOMINAL EXTENSOR MUSCLE.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 8
FT SEQUENCE 8 AA; 995 MW; C6D40729C4576AB5 CRC64;

Query Match 27.8%; Score 10; DB 1; Length 8;
QY 3 NVLQ 6
DB 3 NFLR 6

```

```

Db 3 NFLR 6

RESULT 8
ID FAR2_PROCL STANDARD; PRT; 7 AA.
AC P38498;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Cardioexcitatory FMRFamide homolog DF2
OS Procamburus clarkii (Red swamp crayfish).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
OC Astacidea; Astacoidea; Cambaridae; Procamburus.
OX NCBI_TaxID=6728;
RN [1]
RP SEQUENCE.
RC TISSUE=Pericardial organs;
RX MEDLINE=93248032; PubMed=8387183;
RA Mercier A.J., Orchard I., Tebrugge V., Skerrett M.;
RT "Isolation of two FMRFamide-related peptides from crayfish
RT pericardial organs.";
RL Peptides 14:137-143(1993)
CC -!- FUNCTION: INCREASES THE RATE AND AMPLITUDE OF SPONTANEOUS
CC CONTRACTIONS OF SEMI-ISOLATED HEARTS. INCREASES THE AMPLITUDE OF
CC EXCITATORY POSTSYNAPTIC POTENTIALS IN ABDOMINAL EXTENSOR MUSCLE.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 7
FT SEQUENCE 7 AA; 967 MW; 69D40729C4540AC0 CRC64;

Query Match 27.8%; Score 10; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.1e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 NVLQ 6
DB 3 NFLR 6

RESULT 9
ID FAR1_PANRE STANDARD; PRT; 8 AA.
AC P41872;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FMRFamide-like neuropeptide PFI (SDPNFLRP-amide).
OS Panagrellus redivivus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
OC Panagrolaimoidea; Panagrolaimidae; Panagrellus.
OX NCBI_TaxID=6233;
RN [1]
RP SEQUENCE.
RX MEDLINE=93027659; PubMed=1408999;
RA Geary T.G., Price D.A., Bowman J.W., Winterrowd C.A., Mackenzie C.D.,
RA Garrison R.D., Williams J.F., Friedman A.R.;
RT "Two FMRFamide-like peptides from the free-living nematode
RT Panagrellus redivivus.";
RL Peptides 13:209-214(1992).
CC -!- FUNCTION: MYOACTIVE.
CC -!- TISSUE SPECIFICITY: NERVE CORDS AND PAIRED GROUPS OF CELLS LOCATED
CC CAUDALLY TO THE BASE OF THE PHARYNX.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 8
FT SEQUENCE 8 AA; 995 MW; C6D40729C4576AB5 CRC64;

Query Match 27.8%; Score 10; DB 1; Length 8;
QY 3 NVLQ 6
DB 3 NFLR 6

```

```

Best Local Similarity 50.0%; Pred. No. 1.1e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 NVLQ 6
   | | :
Db 4 NFLR 7

RESULT 10
FARB_HOMAM
ID FARB_HOMAM STANDARD; PRT; 8 AA.
AC P41866;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FMRamide-like neuroptide 3 (FLI 3) (F2).
OS Homarus americanus (American lobster).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
OC Astacidea; Nephropoidea; Nephropidae; Homarus.
OX NCBI_TaxID=6706;
RN [1]
RP SEQUENCE.
RC TISSUE=Pericardial organs;
RX MEDLINE=88116164; PubMed=3429714;
RA Trimmer B.A., Kobierski L.A., Kravitz E.A.;
RT "Purification and characterization of FMRamide-like immunoreactive
RT substances from the lobster nervous system: isolation and sequence
RT analysis of two closely related peptides.";
RL J. Comp. Neurol. 266:16-26(1987).
CC -!- MISCELLANEOUS: PERICARDIAL ORGANS RELEASE THIS PEPTIDE WITH 100 MM
CC POTASSIUM IN THE PRESENCE OF CALCIUM.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuroptide; Amidation.
FT MOD_RES 8 8
SQ SEQUENCE 8 AA; 1054 MW; C6D40729C4540AB5 CRC64;

Query Match 27.8%; Score 10; DB 1; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.1e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 NVLQ 6
   | | :
Db 4 NFLR 7

RESULT 11
FARB_CALVO
ID FARB_CALVO STANDARD; PRT; 7 AA.
AC P41866;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Callifwrfamide 11.
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Oestroidea; Calliphoridae; Calliphora.
OX NCBI_TaxID=27454;
RN [1]
RP SEQUENCE.
RC TISSUE=Thoracic ganglion;
RX MEDLINE=92196111; PubMed=1549595;
RA Duve H., Johnson A.H., Sewell J.C., Scott A.G., Orchard I.,
RA Rehfeld J.F., Thorpe A.;
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT neuroptides (designated callifwrfamides) from the blowfly
RT Calliphora vomitoria.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
PR PIR; B44787; B44787.

KW Neuroptide; Amidation.
FT MOD_RES 7 7
SQ SEQUENCE 7 AA; 926 MW; 69D40699C44AB700 CRC64;

Query Match 25.0%; Score 9; DB 1; Length 7;
Best Local Similarity 20.0%; Pred. No. 1.1e+05;
Matches 1; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 NVLQ 6
   | | :
Db 2 DNEMR 6

RESULT 12
IGAO_DACDE
ID IGAO_DACDE STANDARD; PRT; 7 AA.
AC P06294;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Galactose oxidase inhibitor.
OS Dactylium dendroides (Cladobotryum dendroides).
OC Eukaryota; Fungi; Ascomycota; Pezilomycotina; Sordariomycetes;
OC Hypocreales; Hypocreaceae; Hypomyces.
OX NCBI_TaxID=5132;
RN [1]
RP SEQUENCE.
RA Avigad G., Markus Z.;
RT "Identification of a peptide inhibitor of galactose oxidase from
RT Dactylium dendroides.";
RL Fed. Proc. 31:447-447(1972).
CC -!- FUNCTION: BINDS ONE COPPER ION PER MOLECULE BUT DOES NOT BIND THE
CC GALACTOSE OXIDASE APOENZYME. IT MAY INACTIVATE THE ENZYME BY
CC BINDING TO ITS PROSTHETIC COPPER GROUP.
DR PIR; A01341; XEYDGD.
KW Copper; Metalloenzyme inhibitor.
SQ SEQUENCE 7 AA; 706 MW; 75BB01A456D87DB0 CRC64;

Query Match 25.0%; Score 9; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 OXT 8
   | | :
Db 3 QNT 5

RESULT 13
AKH_TABAT
ID AKH_TABAT STANDARD; PRT; 8 AA.
AC P14595;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Adipokinetic hormone (AKH) (Dipteran corpora cardiaca factor I)
DE (DCC 1).
OS Tabanus atratus (Horse fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Tabanomorpha; Tabanidae; Tabanus.
OX NCBI_TaxID=7207;
RN [1]
RP SEQUENCE.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=90046758; PubMed=2813385;
RA Jarfe H., Raina A.K., Riley C.T., Fraser B.A., Nachman R.J.,
RA Vogel V.W., Zhang Y.-S., Hayes D.K.;
RT "Primary structure of two neuroptide hormones with adipokinetic and
RT hypotrehalosemic activity isolated from the corpora cardiaca of horse
RT flies (Diptera).";
RL Proc. Natl. Acad. Sci. U.S.A. 86:8161-8164(1989).
CC -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF

```



```

CC DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR PTR: A33995; A33995.
DR InterPro: IPR002047; AKH.
DR PROSITE: PS00256; AKH; 1.
KW Neuropeptide; Amidation; Flight.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 949 MW; 86786771A9DIA736 CRC64;

Query Match 25.0%; Score 9; DB 1; Length 8;
Best Local Similarity 66.7%; Pred. No. 1.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 QXT 8
DB 1 QLT 3

RESULT 14
B44K_PORGI STANDARD; PRT; 8 AA.
AC P81886;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 44 kDa immunogenic protein (Fragment).
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides; Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP SEQUENCE.
RC STRAIN=VPB 3492;
RA MEDLINE=20198497; PubMed=10731616;
RA Norris J.M., Love D.N.;
RT "Serum antibody responses of cats to soluble whole cell antigens of
RT feline Porphyromonas gingivalis.";
RL Vet. Microbiol. 73:37-49(2000).
KW Antigen.
CC -!- SIMILARITY: TO P.GINGIVALIS HEMAGGLUTININ A.
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 989 MW; 9554540326CB476D CRC64;

Query Match 25.0%; Score 9; DB 1; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.1e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NV 4
DB 7 NI 8

RESULT 15
COXG_RAT STANDARD; PRT; 8 AA.
AC P80430;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytochrome c oxidase polypeptide VIB (EC 1.9.3.1) (AED) (Fragment).
GN COX6B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RC STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=95324529; PubMed=7601105;
RA Schaeffer H., Noack H., Halangk W., Brandt U., von Jagow G.;
RT "Cytochrome-c oxidase in developing rat heart. Enzymic properties and

```

```

RT amino-terminal sequences suggest identity of the fetal heart and the
RT adult liver isoform."
RL Eur. J. Biochem. 230:235-241(1995).
CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC MITOCHONDRIAL ELECTRON TRANSPORT. THIS PROTEIN MAY BE ONE OF THE
CC HEME-BINDING SUBUNITS OF THE OXIDASE.
CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIB FAMILY.
KW Oxidoreductase; Mitochondrion.
FT NON_TER 1 1
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 1039 MW; 8101E9CAA73AE456 CRC64;

Query Match 25.0%; Score 9; DB 1; Length 8;
Best Local Similarity 66.7%; Pred. No. 1.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 NVL 5
DB 2 NXL 4

Search completed: February 26, 2003, 15:43:43
Job time : 12 secs

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:41:42 : Search time 91 Seconds  
(without alignments)  
18,114 Million cell updates/sec

Title: US-09-476-485A-24  
Perfect score: 36  
Sequence: 1 TNNVLQXT 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 390

Minimum DB seq length: 0  
Maximum DB seq length: 8

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | % Query Match | Length | DB ID     | Description        |
|------------|-------|---------------|--------|-----------|--------------------|
| 1          | 18    | 50.0          | 8      | 10 Q9SB24 | Q9sb24 nicotiana t |
| 2          | 17    | 47.2          | 8      | 2 O87471  | O87471 haemophilus |
| 3          | 15    | 41.7          | 4      | 11 Q08433 | Q08433 rattus norv |
| 4          | 15    | 41.7          | 8      | 4 Q15893  | Q15893 homo sapien |
| 5          | 14    | 38.9          | 8      | 2 Q8RJ10  | Q8rj10 streptomyce |
| 6          | 14    | 38.9          | 8      | 10 Q40659 | Q40659 oryza sativ |
| 7          | 13    | 36.1          | 8      | 3 Q05403  | Q05403 saccharomyc |
| 8          | 13    | 36.1          | 8      | 4 Q8TF70  | Q8tf70 homo sapien |
| 9          | 13    | 36.1          | 8      | 11 Q8R5M9 | Q8r5m9 mus musculu |
| 10         | 12    | 33.3          | 8      | 2 Q45615  | Q45615 bacillus su |
| 11         | 12    | 33.3          | 8      | 2 Q49534  | Q49534 mycoplasma  |
| 12         | 11    | 30.6          | 7      | 2 Q47505  | Q47505 escherichia |
| 13         | 11    | 30.6          | 8      | 4 Q16468  | Q16468 homo sapien |
| 14         | 11    | 30.6          | 8      | 6 Q8WNS1  | Q8wns1 bos taurus  |
| 15         | 11    | 30.6          | 8      | 8 Q9MSX1  | Q9msx1 jurinea hum |
| 16         | 11    | 30.6          | 8      | 11 Q62933 | Q62933 rattus norv |

|    |    |      |   |           |                     |
|----|----|------|---|-----------|---------------------|
| 17 | 11 | 30.6 | 8 | 12 Q9IU21 | Q9iu21 influenza a  |
| 18 | 11 | 30.6 | 8 | 12 Q9IU19 | Q9iu19 influenza a  |
| 19 | 11 | 30.6 | 8 | 13 P87488 | P87488 oncorhynch   |
| 20 | 10 | 27.8 | 7 | 5 P83274  | P83274 macrobrachi  |
| 21 | 10 | 27.8 | 7 | 12 Q9YQ10 | Q9yq10 transmissib  |
| 22 | 10 | 27.8 | 8 | 2 Q9R9C2  | Q9r9c2 borrelia bu  |
| 23 | 10 | 27.8 | 8 | 2 Q93SR0  | Q93sr0 staphylococ  |
| 24 | 10 | 27.8 | 8 | 2 Q53790  | Q53790 streptococc  |
| 25 | 10 | 27.8 | 8 | 4 Q9UMC7  | Q9umc7 homo sapien  |
| 26 | 10 | 27.8 | 8 | 5 Q9UBI3  | Q9ub13 albinaria h  |
| 27 | 10 | 27.8 | 8 | 5 P83275  | P83275 macrobrachi  |
| 28 | 10 | 27.8 | 8 | 5 P83316  | P83316 penaeus mon  |
| 29 | 10 | 27.8 | 8 | 8 Q9T2W0  | Q9t2w0 saccharomyc  |
| 30 | 9  | 25.0 | 7 | 12 Q66113 | Q66113 cherry leaf  |
| 31 | 9  | 25.0 | 7 | 12 Q67113 | Q67113 influenzavi  |
| 32 | 9  | 25.0 | 8 | 2 Q56429  | Q56429 thermus the  |
| 33 | 9  | 25.0 | 8 | 2 P72279  | P72279 rhodococcus  |
| 34 | 9  | 25.0 | 8 | 2 Q51594  | Q51594 escherichia  |
| 35 | 9  | 25.0 | 8 | 2 P77556  | P77556 escherichia  |
| 36 | 9  | 25.0 | 8 | 4 Q15894  | Q15894 homo sapien  |
| 37 | 9  | 25.0 | 8 | 4 Q15901  | Q15901 homo sapien  |
| 38 | 9  | 25.0 | 8 | 4 Q9UDZ4  | Q9udz4 homo sapien  |
| 39 | 9  | 25.0 | 8 | 4 Q96QF9  | Q96qf9 homo sapien  |
| 40 | 9  | 25.0 | 8 | 8 Q9GD47  | Q9gd47 hydrastelle  |
| 41 | 9  | 25.0 | 8 | 8 Q9GCZ4  | Q9gcz4 nenga punil  |
| 42 | 9  | 25.0 | 8 | 8 Q94VP6  | Q94vf6 varanus job  |
| 43 | 9  | 25.0 | 8 | 8 Q8WGC9  | Q8wgc9 upogebia af  |
| 44 | 9  | 25.0 | 8 | 8 Q8WGC7  | Q8wgc7 petrollisthe |
| 45 | 9  | 25.0 | 8 | 10 Q9SAY7 | Q9say7 dioscorea t  |

## ALIGNMENTS

|                       |                                                                     |                    |        |            |    |
|-----------------------|---------------------------------------------------------------------|--------------------|--------|------------|----|
| RESULT 1              |                                                                     |                    |        |            |    |
| Q9SB24                |                                                                     |                    |        |            |    |
| ID                    | Q9SB24                                                              | PRELIMINARY;       | PRT;   | 8 AA.      |    |
| AC                    | Q9SB24;                                                             |                    |        |            |    |
| DT                    | 01-MAY-2000 (TREMBlrel. 13, Created)                                |                    |        |            |    |
| DT                    | 01-MAY-2000 (TREMBlrel. 13, Last sequence update)                   |                    |        |            |    |
| DE                    | 01-JUN-2001 (TREMBlrel. 17, Last annotation update)                 |                    |        |            |    |
| DE                    | ParAt protein (Fragment).                                           |                    |        |            |    |
| OS                    | Nicotiana tabacum (Common tobacco).                                 |                    |        |            |    |
| OC                    | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  |                    |        |            |    |
| OC                    | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;        |                    |        |            |    |
| OC                    | Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.          |                    |        |            |    |
| OX                    | NCBI_TaxID=4097;                                                    |                    |        |            |    |
| RN                    | [1]                                                                 |                    |        |            |    |
| RP                    | SEQUENCE FROM N.A.                                                  |                    |        |            |    |
| RA                    | Niwa Y., Muranaka T., Baba A., Machida Y.;                          |                    |        |            |    |
| RT                    | "Organ-specific and auxin-inducible expression of two tobacco para- |                    |        |            |    |
| RT                    | related genes in transgenic plants.";                               |                    |        |            |    |
| RL                    | DNA Res. 0:0-0(1994).                                               |                    |        |            |    |
| DR                    | EMBL; D42119; BAA07700.1; -.                                        |                    |        |            |    |
| FT                    | NON_TER 8                                                           |                    |        |            |    |
| SQ                    | SEQUENCE 8 AA; 905 MW; FE32D2C44455BB16 CRC64;                      |                    |        |            |    |
| Query Match           | 50.0%;                                                              | Score 18;          | DB 10; | Length 8;  |    |
| Best Local Similarity | 60.0%;                                                              | Pred. No. 6.7e+05; |        |            |    |
| Matches               | 3;                                                                  | Conservative       | 2;     | Mismatches | 0; |
|                       |                                                                     |                    |        | Indels     | 0; |
|                       |                                                                     |                    |        | Gaps       | 0; |
| QY                    | 1 TNNVL 5                                                           |                    |        |            |    |
|                       |                                                                     |                    |        |            |    |
| Db                    | 3 SNNVV 7                                                           |                    |        |            |    |
| RESULT 2              |                                                                     |                    |        |            |    |
| O87471                |                                                                     |                    |        |            |    |
| ID                    | O87471                                                              | PRELIMINARY;       | PRT;   | 8 AA.      |    |
| AC                    | O87471;                                                             |                    |        |            |    |
| DT                    | 01-NOV-1998 (TREMBlrel. 08, Created)                                |                    |        |            |    |
| DT                    | 01-NOV-1998 (TREMBlrel. 08, Last sequence update)                   |                    |        |            |    |
| DT                    | 01-NOV-1998 (TREMBlrel. 08, Last annotation update)                 |                    |        |            |    |

```

DE Hifa (Fragment).
GN HIFA.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EAGAN;
EX MEDLINE=98389689; PubMed=9721313;
RA Mhlanga-Mutagadura T., Morlin G., Smith A.L., Eisenstark A.,
RA Golomb M.;
RT "Evolution of the major pilus gene cluster of haemophilus
RT influenzae.";
RL J. Bacteriol. 180:4693-4703(1998).
DR EMBL: AF071762; AAC35830.1; -.
FT NON_TER 1 1
FT SEQUENCE 8 AA; 876 MW; DAB44451A7272325 CRC64;
SQ SEQUENCE 8 AA; 876 MW; DAB44451A7272325 CRC64;

Query Match 47.2%; Score 17; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNN 3
Db 5 TNN 7

RESULT 3
Q08433 PRELIMINARY; PRT; 4 AA.
ID Q08433
AC Q08433;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE UDP-glucuronosyltransferase, microsomal (EC 2.4.1.17) (UDPGT)
DE (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GUNN;
EX MEDLINE=91282758; PubMed=1840486;
RA Sato H., Aono S., Kashiwamata S., Koiwai O.;
RT "Genetic defect of bilirubin UDP-glucuronosyltransferase in the
RT hyperbilirubinemic Gunn rat.";
RL Biochem. Biophys. Res. Commun. 177:1161-1164(1991).
CC -!- FUNCTION: UDPGT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND
CC SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND
CC ENDOGENOUS COMPOUNDS.
CC -!- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR
CC -!- BETA-D-GLUCURONOSIDE.
CC EMBL: S38636; AAB19259.1; -.
KW transferase; Glycosyltransferase; Microsome; Multigene family.
FT NON_TER 1 1
FT NON_TER 4 4
FT SEQUENCE 4 AA; 473 MW; 633732C420000000 CRC64;
SQ SEQUENCE 4 AA; 473 MW; 633732C420000000 CRC64;

Query Match 41.7%; Score 15; DB 11; Length 4;
Best Local Similarity 75.0%; Pred. No. 6.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 NVLQ 6
Db 1 NVLK 4

RESULT 4
Q15893 PRELIMINARY; PRT; 8 AA.
ID Q15893

```

```

AC Q15893;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE (clone XP587A) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of
RT arrayed cDNAs and cosmid libraries.";
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL: L32073; AAA73883.1; -.
FT NON_TER 1 1
FT NON_TER 8 8
FT SEQUENCE 8 AA; 874 MW; DAA1B6D7376456C5 CRC64;
SQ SEQUENCE 8 AA; 874 MW; DAA1B6D7376456C5 CRC64;

Query Match 41.7%; Score 15; DB 4; Length 8;
Best Local Similarity 37.5%; Pred. No. 6.7e+05;
Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TNNVLOXT 8
Db 1 SQNPLOYS 8

RESULT 5
Q8RJ10 PRELIMINARY; PRT; 8 AA.
ID Q8RJ10
AC Q8RJ10;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE DNA replication initiation protein (Fragment).
GN DNAA-LIKE.
OS Streptomyces caespitosus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=53502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC27422;
RA Ma W., Mao X., Jie L., Qin Z.J., Jiang W.H., Chiao R.S., Zhao G.P.;
RT "Cloning, sequence analysis and function analysis of the replication
RT origin oric of Streptovorticillum caespitosus ATCC27422.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ458440; CAD30324.1; -.
FT NON_TER 1 1
FT SEQUENCE 8 AA; 915 MW; 04744330440451A7 CRC64;
SQ SEQUENCE 8 AA; 915 MW; 04744330440451A7 CRC64;

Query Match 38.9%; Score 14; DB 2; Length 8;
Best Local Similarity 50.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNV 4
Db 2 TNRI 5

RESULT 6
Q40659 PRELIMINARY; PRT; 8 AA.
ID Q40659
AC Q40659;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Alpha-amylase (Fragment).

```

```

OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91078641; PubMed=2258052;
RA Kumagai M.H., Shah M., Terashima M., Vrkljan Z., Whitaker J.R.,
RA Rodriguez R.L.;
RT "Expression and secretion of rice alpha-amylase by saccharomycers
RT cerevisiae.";
RL Gene 94:209-216(1990).
DR EMBL; M62916; AAA33892.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 948 MW; EBC694444732D6D6 CRC64;

Query Match 38.9%; Score 14; DB 10; Length 8;
Best Local Similarity 50.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNVL 5
Db 5 NNWV 8

RESULT 7
Q05403 PRELIMINARY; PRT; 8 AA.
AC Q05403;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE DNA for ORF's from chromosome XV (Fragment).
GN COQ3 AND YOL096C.
OS Saccharomycetes cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=FV1679;
RX MEDLINE=96021609; PubMed=8533473;
RA Zumslein E., Pearson B.M., Kalogeropoulos A., Schweizer M.;
RT "A 29,425 kb segment on the left arm of yeast chromosome XV contains
RT more than twice as many unknown as known open reading frames.";
RL Yeast 11:975-986(1995).
DR EMBL; X83121; CAA58183.1; -.
DR SGD; S0005456; COQ3.
FT NON_TER 8
SQ SEQUENCE 8 AA; 879 MW; 7B5322D2C441E058 CRC64;

Query Match 36.1%; Score 13; DB 3; Length 8;
Best Local Similarity 40.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNVLQ 6
Db 3 HNVVK 7

RESULT 8
Q8TF70 PRELIMINARY; PRT; 8 AA.
AC Q8TF70;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Microphthalma-associated transcription factor (Fragment).
GN MITF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Takeda K., Yasumoto K., Shibahara S.;
RT "An Mitf isoform exclusively expressed in the affected cells of Mitf
RT mutants.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB061771; BAB85121.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 989 MW; ED0727204415A1A6 CRC64;

Query Match 36.1%; Score 13; DB 4; Length 8;
Best Local Similarity 40.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVL 5
Db 2 TSRIL 6

RESULT 9
Q8R5M9 PRELIMINARY; PRT; 8 AA.
AC Q8R5M9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Microphthalma-associated transcription factor (Fragment).
GN MITF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/C; TISSUE=EYE;
RA Takeda K., Yasumoto K., Shibahara S.;
RT "An Mitf isoform exclusively expressed in the affected cells of Mitf
RT mutants.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB061769; BAB85120.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 989 MW; ED0727204415A1A6 CRC64;

Query Match 36.1%; Score 13; DB 11; Length 8;
Best Local Similarity 40.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVL 5
Db 2 TSRIL 6

RESULT 10
Q45615 PRELIMINARY; PRT; 8 AA.
AC Q45615;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GutB protein (Fragment).
GN GUTB.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RX MEDLINE=94253000; PubMed=8195086;
RA Ye R., Wong S.L.;
RT "Transcriptional regulation of the Bacillus subtilis glucitol
RT dehydrogenase gene.";

```

RL J. Bacteriol. 176:3314-3320(1994).  
 DR EMBL; L16626; AAA20875.1; -;  
 FT NON\_TER 8  
 SQ SEQUENCE 8 AA; 927 MW; FD56C772D1A1F1A6 CRC64;

Query Match 33.3%; Score 12; DB 2; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 6.7e+05;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
 I: | |  
 Db 2 THTVPQ 7

RESULT 11  
 Q49534 PRELIMINARY; PRT; 8 AA.  
 ID Q49534  
 AC Q49534;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE P120 (Fragment).  
 GN P120.  
 OS Mycoplasma hominis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 OC Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2098;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=V2785;  
 RA Nyvold C., Birkelund S., Christiansen G.;  
 RT "The Mycoplasma hominis P120 membrane protein gene contains a 659 base pair hypervariable domain."  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U22025; AAA67455.1; -;  
 FT NON\_TER 1  
 SQ SEQUENCE 8 AA; 869 MW; 914457605B02C05D CRC64;

Query Match 33.3%; Score 12; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 6.7e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NN 3  
 I: | |  
 Db 7 NN 8

RESULT 12  
 Q47505 PRELIMINARY; PRT; 7 AA.  
 ID Q47505  
 AC Q47505;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE MCCA protein.  
 GN MCCA.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96099297; PubMed=8522520;  
 RA Gonzalez-Pastor J.E., San Millan J.L., Castilla M.A., Moreno F.;  
 RT "Structure and organization of plasmid genes required to produce the translation inhibitor microcin C7."  
 RL J. Bacteriol. 177:7131-7140(1995).  
 DR EMBL; X57583; CAA40808.1; -;  
 KW Plasmid.  
 SQ SEQUENCE 7 AA; 763 MW; 644DD44861B406F0 CRC64;

Query Match 30.6%; Score 11; DB 2; Length 7;

Best Local Similarity 66.7%; Pred. No. 6.7e+05;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNN 3  
 I: | |  
 Db 3 TGN 5

RESULT 13  
 Q16468 PRELIMINARY; PRT; 8 AA.  
 ID Q16468  
 AC Q16468;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE DNA for cosmid cC13-1134 PCR primer 1 (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96435920; PubMed=8838806;  
 RA James L.A., Ogilvie D.J., Yamakawa K., Nakamura Y., Stirling C.J.,  
 RA Anand R.;  
 RT "Walking, cloning and mapping with YACs in 3q27. Localisation of 5 RT ESTs including 3 members of the Cystatin gene family and RT identification of CpG islands."  
 RL Genomics 32:425-430(1996).  
 DR EMBL; X88976; CAA61407.1; -;  
 FT NON\_TER 8  
 SQ SEQUENCE 8 AA; 925 MW; FD5411A7376871E6 CRC64;

Query Match 30.6%; Score 11; DB 4; Length 8;  
 Best Local Similarity 66.7%; Pred. No. 6.7e+05;  
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNN 3  
 I: | |  
 Db 6 TRN 8

RESULT 14  
 Q8WNS1 PRELIMINARY; PRT; 8 AA.  
 ID Q8WNS1  
 AC Q8WNS1;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE X-linked zinc finger protein (Fragment).  
 GN ZFX.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Poloumienko A., Blecher S.;  
 RT "Comparison between intron-exon structures in zFX and zFY genes."  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF045782; AAL58190.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 8  
 SQ SEQUENCE 8 AA; 904 MW; DF1DC2C4472AAB1A CRC64;

Query Match 30.6%; Score 11; DB 6; Length 8;  
 Best Local Similarity 50.0%; Pred. No. 6.7e+05;  
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 NVLQ 6  
 I: | |  
 Db 5 NVAE 8

```

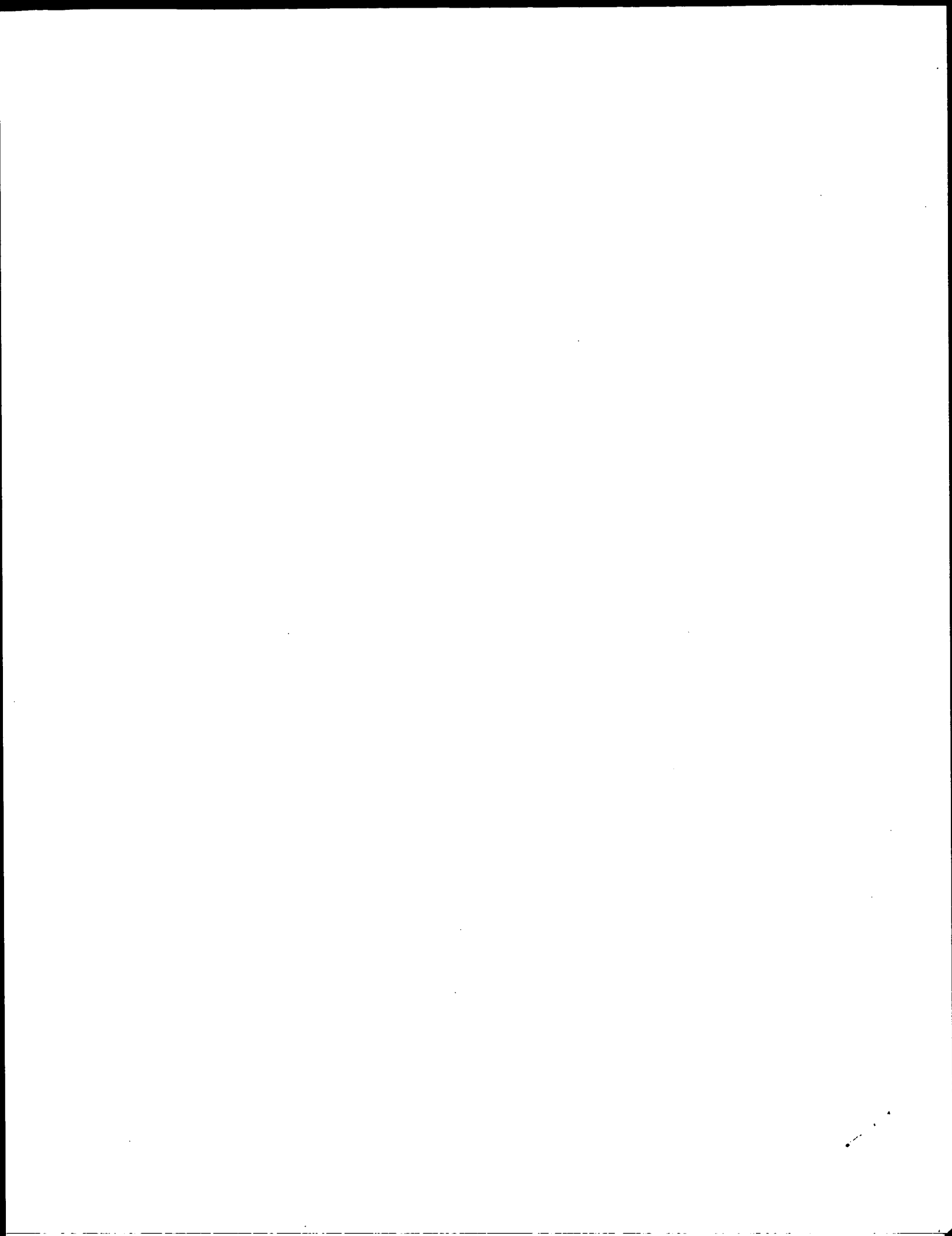
RESULT 15
Q9MSX1
ID Q9MSX1 PRELIMINARY; PRT; 8 AA.
AC Q9MSX1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE PsbA (Fragment).
DE PSBA.
GN Jurinea humilis.
OS Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Asterales; Asteraceae; Cardueae; Jurinea.
OX NCBI_TaxID=41594;
RN [1]
RP SEQUENCE FROM N.A.
RA O'Hanlon P.C., Briese D.T.;
RT "Evidence for hierarchical and non-hierarchical evolution in the
RT Carduinae thistles";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF129846; AAF78138.1; -
KW Chloroplast.
FT NON_TER
SQ SEQUENCE 8 AA; 788 MW; 457451B5A76DDB10 CRC64;

Query Match 30.6%; Score 11; DB 8; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TN 2
   ||
Db 6 TN 7

```

Search completed: February 26, 2003, 15:45:22  
 Job time : 93 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:38:41 ; Search time 34 seconds  
(without alignments)  
31.353 Million cell updates/sec

Title: US-09-476-485A-24

Perfect score: 36  
Sequence: 1 TNNVLQXT 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 86491

Minimum DB seq length: 0  
Maximum DB seq length: 8

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

```

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length DB | ID          | Description        |
|------------|-------|-------------|-----------|-------------|--------------------|
| 1          | 34    | 94.4        | 8         | 20 AAW87974 | Lectin derived pro |
| 2          | 34    | 94.4        | 8         | 22 AAG62889 | Peptide derived fr |
| 3          | 22    | 61.1        | 6         | 20 AAY23824 | Peptide derived fr |
| 4          | 20    | 55.6        | 7         | 18 AAW27518 | Human interleukin  |
| 5          | 20    | 55.6        | 8         | 15 AAR57168 | Peptide corresp. t |
| 6          | 19    | 52.8        | 7         | 21 AAB23214 | Hsp47-binding phag |
| 7          | 19    | 52.8        | 7         | 22 AAB72780 | Antibacterial pept |
| 8          | 19    | 52.8        | 7         | 22 AAB72821 | Antibacterial pept |
| 9          | 19    | 52.8        | 7         | 22 AAB72823 | Antibacterial pept |
| 10         | 18    | 50.0        | 5         | 15 AAR54171 | CHA255 light chain |

|    |    |      |   |             |                    |
|----|----|------|---|-------------|--------------------|
| 11 | 18 | 50.0 | 5 | 21 AAB52719 | T21/DP107 peptide  |
| 12 | 18 | 50.0 | 5 | 21 AAB14673 | HIV-1 isolate LAI  |
| 13 | 18 | 50.0 | 5 | 22 AAB54937 | Anti-HIV peptide D |
| 14 | 18 | 50.0 | 6 | 20 AAY23409 | V beta 6 clone fou |
| 15 | 18 | 50.0 | 6 | 20 AAW78331 | Human fibrinogen g |
| 16 | 18 | 50.0 | 6 | 21 AAB52720 | T21/DP107 peptide  |
| 17 | 18 | 50.0 | 6 | 21 AAB14674 | HIV-1 isolate LAI  |
| 18 | 18 | 50.0 | 6 | 22 AAB54938 | Anti-HIV peptide D |
| 19 | 18 | 50.0 | 7 | 21 AAB52721 | T21/DP107 peptide  |
| 20 | 18 | 50.0 | 7 | 21 AAB14675 | HIV-1 isolate LAI  |
| 21 | 18 | 50.0 | 7 | 21 AAY81478 | Bovine membrane-bo |
| 22 | 18 | 50.0 | 7 | 22 AAB72812 | Antibacterial pept |
| 23 | 18 | 50.0 | 7 | 22 AAB54939 | Anti-HIV peptide D |
| 24 | 18 | 50.0 | 8 | 21 AAB52722 | T21/DP107 peptide  |
| 25 | 18 | 50.0 | 8 | 21 AAB14676 | HIV-1 isolate LAI  |
| 26 | 18 | 50.0 | 8 | 22 AAB12040 | HIV A02 super moti |
| 27 | 18 | 50.0 | 8 | 22 AAB54940 | Anti-HIV peptide D |
| 28 | 17 | 47.2 | 4 | 21 AAB52718 | T20/DP178 peptide  |
| 29 | 17 | 47.2 | 4 | 21 AAB14672 | HIV-1 isolate LAI  |
| 30 | 17 | 47.2 | 4 | 22 AAB54936 | Anti-HIV peptide D |
| 31 | 17 | 47.2 | 5 | 15 AAR54176 | CHA255 light chain |
| 32 | 17 | 47.2 | 5 | 15 AAR54167 | CHA255 light chain |
| 33 | 17 | 47.2 | 5 | 15 AAR54173 | CHA255 light chain |
| 34 | 17 | 47.2 | 5 | 15 AAR54175 | CHA255 light chain |
| 35 | 17 | 47.2 | 5 | 19 AAW65349 | Peptide #10 useful |
| 36 | 17 | 47.2 | 5 | 19 AAW65350 | Peptide #11 useful |
| 37 | 17 | 47.2 | 5 | 19 AAW65351 | Peptide #12 useful |
| 38 | 17 | 47.2 | 5 | 21 AAB37736 | Fibronectin-derive |
| 39 | 17 | 47.2 | 5 | 21 AAB37737 | Fibronectin-derive |
| 40 | 17 | 47.2 | 5 | 21 AAB37738 | Fibronectin-derive |
| 41 | 17 | 47.2 | 5 | 21 AAY80733 | Fibroblast invasio |
| 42 | 17 | 47.2 | 5 | 21 AAY80734 | Fibroblast invasio |
| 43 | 17 | 47.2 | 5 | 21 AAY80735 | Fibroblast invasio |
| 44 | 17 | 47.2 | 5 | 22 AAB68154 | Meat protein deriv |
| 45 | 17 | 47.2 | 5 | 22 AAB37554 | CD8 activation com |

#### ALIGNMENTS

##### RESULT 1

AAW87974

ID AAW87974 standard; Peptide; 8 AA.

AC AAW87974;

DT 13-APR-1999 (first entry)

DE Lectin derived progenitor cell preservation factor derived peptide.

KW Lectin derived progenitor cell preservation factor; progenitor cell;

KW haematopoietic cell; cultured cell preservation; anticancer therapy;

KW myeloablative therapy; sickle-cell anaemia; ablative therapy protection;

KW FLK2/FLT3 receptor; ss.

OS Dolichos lab lab.

PN WO9859038-A1.

PD 30-DEC-1998.

XX 23-JUN-1998; 98WO-US13046.

XX 24-JUN-1997; 97US-0881189.

XX (IMCL-) IMCLONE SYSTEMS INC.

XX (REGC) UNIV CALIFORNIA.

PI Chrispeels MJ, Colucci MG, Moore JG;

DR WPI; 1999-081274/07.

PT New nucleic acid encoding plant lectin that preserves progenitor

PT cells - particularly haematopoietic progenitors, useful for bone  
 PT marrow reconstitution after ablative therapy, and to increase DNA  
 PT transfer in gene therapy  
 PS Claim 1; Page 46; 72pp; English.  
 XX  
 CC The present sequence represents a peptide of lectin derived progenitor  
 CC a cell preservation factor. The protein is used to preserve unipotent,  
 CC pluripotent or totipotent progenitor cells, especially haematopoietic  
 CC cells, and also progenitors from nerve, muscle, skin, gut, bone,  
 CC kidney, liver, pancreas or thymus. Specific applications are  
 CC preservation of cultured cells intended for administration after  
 CC (anticancer) myeloablative therapy (bone marrow or whole-body irradiation  
 CC or chemotherapy) to reconstitute the haematopoietic system; enrichment  
 CC of progenitor cells (e.g. during ex vivo purging of malignant cells);  
 CC treatment of tissues containing haematopoietic progenitors for subsequent  
 CC transplant to improve haematopoietic competence; improving transfer of  
 CC exogenous DNA to progenitor cells (in gene therapy of various  
 CC haematological disorders, e.g. sickle-cell anaemia); and protection  
 CC against ablative therapy (to eliminate proliferating cells specifically),  
 CC followed by re-establishment of differentiation and proliferation of  
 CC preserved progenitors. The protein, when linked to magnetic beads, may  
 CC also be used to isolate cells that express the FLK2/FLT3 receptor.  
 XX  
 SQ Sequence 8 AA;

Query Match 94.4%; Score 34; DB 20; Length 8;  
 Best Local Similarity 87.5%; Pred. No. 7.8e+05;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8  
 |||||  
 Db 1 TNNVLQXT 8

RESULT 2  
 AAG62889  
 ID AAG62889 standard; peptide; 8 AA.  
 XX  
 AC AAG62889;

DT 17-SEP-2001 (first entry)

DE Peptide derived from a hyacinth bean FRIL polypeptide.

XX FRIL; FLK2/Flt3 tyrosine kinase receptor-interacting lectin;  
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;  
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
 KW severe combined immunodeficiency; aplastic anemia; tissue repair.  
 XX  
 OS Dolichos lab lab.

XX Key Location/Qualifiers  
 FH Misc-difference 7  
 FT /note= "not specified"

XX WO200149851-A1.

XX 12-JUL-2001.

XX 30-DEC-1999; 99WO-US31307.

XX 30-DEC-1999; 99WO-US31307.

XX (PHYL-) PHYLOGIX LLC.

XX Colucci MG, Chrispeels MJ, Moore JG;

XX WPI; 2001-441882/47.

XX Legume Progenitor cell preservation factors for in vivo or ex vivo  
 PT preservation of haematopoietic progenitor cells and as therapeutics for  
 PT alleviating/reducing progenitor cell-depleting activity of cancer

PT therapeutics -  
 XX  
 PS Disclosure; Page 19; 173pp; English.  
 XX

CC The present sequence is derived from a FRIL (FLK2/Flt3 tyrosine kinase  
 CC receptor-interacting lectin) polypeptide. The specification describes a  
 CC composition of one or more members of FRIL family of progenitor cell  
 CC preservation factors. The composition is useful for alleviating or  
 CC reducing the hematopoietic progenitor cell-depleting activity of a  
 CC therapeutic treatment, including radiotherapeutic and/or  
 CC chemotherapeutic treatments. Administration of FRIL compositions to a  
 CC patient prior to treatment of the patient with a therapeutic treatment  
 CC having a hematopoietic progenitor cell-depleting activity alleviates or  
 CC reduces the hematopoietic progenitor cell-depleting activity of the  
 CC therapeutic treatment in the patient. FRIL family members are useful for  
 CC isolating population of progenitor cells, hemangioblasts, and mesenchymal  
 CC stem cells. The composition is administered to reduce progenitor cell  
 CC depleting effects of chemotherapeutics, so that the patient can receive  
 CC a higher dose of the chemotherapeutic and preferably recover from cancer.  
 CC It is also administered to patients having, or predisposed to developing  
 CC a condition where the patients hematopoietic progenitor cells are  
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
 CC The isolated mesenchymal cells are useful for tissue repair.  
 XX

SQ Sequence 8 AA;

Query Match 94.4%; Score 34; DB 22; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 7.9e+05;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8  
 |||||  
 Db 1 TNNVLQXT 8

RESULT 3  
 AAY23824

ID AAY23824 standard; Peptide; 6 AA.

XX AAY23824;

DT 14-SEP-1999 (first entry)

DE Peptide derived from gyrase protein and used to design primers.

XX Identification; detection; microbe; gyrase gene; gyrase protein.

XX Unidentified.

XX JF11169175-A.

XX 29-JUN-1999.

XX 12-DEC-1997; 97JP-0343316.

XX 12-DEC-1997; 97JP-0343316.

XX (MARI-) MARINE BIOTECHNOLOGY INST CO LTD.

XX (KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.

XX WPI; 1999-422615/36.

PT Identification and detection of a microbe - by detection of a  
 PT gyrase gene

XX Example 2; Page 36; 42pp; Japanese.

CC The specification describes a method for the identification or  
 CC detection of a microbe, using the gyrase gene as the index. The  
 CC method involves the use of PCR primers to amplify DNA from the  
 CC microbe, which is then identified or detected depending on its  
 CC base sequence. The method can be used to classify and identify  
 CC an unidentified microbe strain rapidly and with high precision.

CC The present sequence represents a peptide derived from gyrase  
 CC protein and used to design primers for use in the method of the  
 CC invention.  
 XX  
 XX  
 SQ Sequence 6 AA;

Query Match 61.1%; Score 22; DB 20; Length 6;  
 Best Local Similarity 66.7%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
 ||| :  
 Db 1 TNNIPQ 6

RESULT 4  
 AAW27518  
 ID AAW27518 standard; peptide; 7 AA.

XX AC AAW27518;  
 XX AC  
 XX 09-JAN-1998 (first entry)  
 XX XX  
 XX Human interleukin 9 receptor antigenic peptide.

XX Antagonist; human; interleukin 9; asthma associated factor 1; IL-9;  
 KW atopic allergy; asthma; bronchial hypersensitivity; BHR; eczema;  
 KW rhinitis; urticaria; allergic inflammation; bowel; alleviation;  
 KW treatment; antigen; antibody; receptor.

XX Homo sapiens.

XX WO9708321-A1.

XX 06-MAR-1997.

XX 23-AUG-1996; 96WO-US12757.

XX 06-AUG-1996; 96US-0023800.

XX 24-AUG-1995; 95US-0002765.

XX (MAGA-) MAGAININ PHARM INC.

XX Lee MW, Levitt RC, Nicholas N, Prasad KU;

XX WPI; 1997-179278/16.

XX Human interleukin-9 variant with Met at position 117 - useful for  
 XX treating atopic allergy, esp. asthma

XX Claim 105; Page 105; 142pp; English.

XX The present peptide is an antigen derived from human interleukin 9  
 CC (hIL-9) receptor, also known as asthma associated factor 1, which  
 CC can be used to raise antibodies (Ab) that block the binding of  
 CC hIL-9 to its receptor. hIL-9 plays a role in atopic allergy,  
 CC asthma and related disorders, e.g. bronchial hypersensitivity  
 CC (BHR), rhinitis, urticaria, allergic inflammatory disorders of the  
 CC bowel and various forms of eczema. Compounds which down regulate  
 CC the function of IL-9, e.g. antisense DNA, Ab, truncated (soluble)  
 CC IL-9 and its splice variants and other IL-9 or IL-9 receptor  
 CC antagonists, are useful for alleviating such diseases. Anti-IL-9  
 CC Ab are also useful for quantifying IL-9 in a sample.

XX SQ Sequence 7 AA;

Query Match 55.6%; Score 20; DB 18; Length 7;  
 Best Local Similarity 75.0%; Pred. No. 7.8e+05;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4  
 ||| :  
 Db 4 TNNI 7

RESULT 5  
 AAR57168

ID AAR57168 standard; Protein; 8 AA.

XX AC AAR57168;

XX 28-FEB-1995 (first entry)

XX Peptide corresp. to E.coli CS31A subunit residues 88-96.

XX Capsule protein; CS31A; surface antigen; immunogen; vaccine;  
 KW insertion; modification; mutation; epitope; poliovirus;  
 KW porcine transmissible gastroenteritis virus; bovine rotavirus;  
 KW foot and mouth disease virus; clpG gene.

XX Escherichia coli (strain 31A).

XX FR2699538-A.

XX 24-JUN-1994.

XX 22-DEC-1992; 92FR-0015464.

XX 22-DEC-1992; 92FR-0015464.

XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.

XX Bousquet F, Der Vartanian M, Girardeau J, Martin C;  
 PI Mechlin M;

XX WPI; 1994-227158/28.

XX Sub-unit of capsular protein CS31A contg heterologous sequences -  
 PT useful for representation of heterologous epitopes in vaccines,  
 PT for antibody prodn and in immunological tests

XX Disclosure; Page 36; 128pp; French.

XX The CS31A subunit of the E.coli capsular protein can accept large  
 CC inserts (up to 100 amino acids) without inhibition of biosynthesis.  
 CC The insertions are esp. epitope sequences from viruses and the  
 CC modified subunits, or microorganisms expressing them, are useful as  
 CC immunogens. Pref. the subunit has at least one insertion and/or  
 CC substitution by a heterologous peptide in one of the regions -13 to  
 CC +8, 10-58, 123-164 or 183-257. Esp. the modifications are made in  
 CC the regions corresp. to natural epitopes of CS31A as these are the  
 CC most accessible sites on the folded protein. The epitopes were  
 CC identified and localised by assaying antipeptide capture.  
 CC This peptide was shown to correspond to an accessible region.

XX SQ Sequence 8 AA;

Query Match 55.6%; Score 20; DB 15; Length 8;  
 Best Local Similarity 57.1%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TNNVLQXT. 8  
 ||| :  
 Db 2 NGVLQSS 8

RESULT 6  
 AAB23214

ID AAB23214 standard; peptide; 7 AA.

XX AC AAB23214;

XX 29-JAN-2001 (first entry)

XX Hsp47-binding phage display heptapeptide, SEQ ID NO:36.  
 DE  
 XX

KW Hsp47-binding heptapeptide; phage display library;  
 KW Hsp47 external domain; carcinoma; cancer; targeting molecule;  
 KW therapy; diagnosis; detection; imaging; drug delivery; invasion;  
 KW migration; metastasis; modulation; tumour; skin; basal cells; colon;  
 KW large intestine; lung; breast; bladder; oral cancer;  
 KW head and neck cancer; larynx; nasopharynx; adrenal cortex;  
 KW apocrine gland; kidney; liver; pancreas; prostate.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200054805-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000WO-US06588.  
 XX  
 PR 15-MAR-1999; 99US-0124481.  
 XX  
 PA (UYNV-) UNIV MARYLAND BALTIMORE.  
 XX  
 PI Sauk JJ;  
 XX  
 DR WPI; 2000-655997/63.  
 XX  
 PT Treating, diagnosing or modulating a carcinoma cell, which expresses  
 PT Heat shock protein 47 on its surface, involves administering an agent  
 PT comprising targeting moiety which binds to Hsp47 external domain -  
 XX  
 PS Example 5; Page 9; 87pp; English.  
 XX  
 CC The invention relates to methods of treating and diagnosing carcinomas  
 CC in which heat shock protein 47 (Hsp47) is expressed on the surface  
 CC of the carcinoma cells, involving administering an agent comprising a  
 CC targeting moiety which specifically binds to the external domain of  
 CC Hsp47. The invention also relates to peptides (AAB23181-B23203) which  
 CC specifically bind to external domains of such surface-localised Hsp47  
 CC molecules and have sequences encompassed by the generic sequences  
 CC XHVVHXXHXXHXXHXX or HXXHXXHXXHXXHXX where X, independently, can be  
 CC any amino acid and HY, independently, can be any hydrophobic amino acid.  
 CC The invention also encompasses methods of screening for agents which  
 CC bind Hsp47 external domains. Hsp47-binding agents can be used to  
 CC treat Hsp47-expressing carcinomas, and for modulating the activity of a  
 CC tumour cell with respect to invasion, migration, motility or metastasis,  
 CC or to its interaction with the extracellular matrix. The targeting  
 CC moiety (such as an Hsp47-binding peptide, a toxin or an antibody) may be  
 CC coupled with a therapeutic moiety (such as a cytotoxic agent or a  
 CC therapeutic gene) for cancer treatment, or with a detectable moiety for  
 CC imaging. Carcinomas which may be treated or diagnosed according to  
 CC methods of the invention include those of the skin, basal cells, large  
 CC intestine, lung, colon, breast, bladder, oral, head and neck, larynx,  
 CC nasopharynx, adrenal cortex, apocrine glands, kidney, liver, pancreas, or  
 CC prostate. Targeting carcinoma cells with Hsp47-binding agents results in  
 CC efficient delivery of therapeutic agents, reduced doses, reduced side  
 CC effects and sensitive detection or imaging of carcinoma cells. Sequences  
 CC AAB23204-B23240 represent phage display library heptapeptides identified  
 CC as being able to bind Hsp47 in an exemplification of the invention.  
 XX  
 SQ Sequence 7 AA;  
 Query Match 52.8%; Score 19; DB 21; Length 7;  
 Best Local Similarity 80.0%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 TNNVL 5  
 DB 3 TNNVL 7  
 RESULT 7  
 AAB72780  
 ID AAB72780 standard; Peptide; 7 AA.  
 XX  
 AC AAB72780;

XX 09-MAY-2001 (first entry)  
 DT Antibacterial peptide SEQ ID NO: 1.  
 DE  
 XX  
 KW Antibacterial compound; Gram-negative bacterium; pilus; chaperone;  
 KW biofilm; disease treatment; bacterial infection.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200110386-A2.  
 XX  
 PD 15-FEB-2001.  
 XX  
 PF 11-AUG-2000; 2000WO-US22087.  
 XX  
 PR 11-AUG-1999; 99US-0148280.  
 XX  
 PA (UNIW ) UNIV WASHINGTON.  
 XX  
 PI Hultgren SJ, Sauer FG, Waksman G, Fuetterer K;  
 XX  
 DR WPI; 2001-226496/23.  
 XX  
 PT An isolated compound for inhibiting pilus assembly -  
 XX  
 PS Claim 12; Page 130; 144pp; English.  
 XX  
 CC The present invention provides antibacterial compounds which are able to  
 CC interfere with Gram-negative bacteria pilus formation and assembly, and  
 CC pilus interaction with chaperone proteins. These are useful in the  
 CC treatment of bacterial infection, and in the prevention of biofilm  
 CC formation. They are particularly useful against *Escherichia coli*,  
 CC *Haemophilus influenzae*, *Salmonella enteritidis*, *S. typhimurium*,  
 CC *Bordetella pertussis*, *Yersinia enterocolitica*, *Helicobacter pylori* and  
 CC *Klebsiella pneumoniae*.  
 XX  
 SQ Sequence 7 AA;  
 Query Match 52.8%; Score 19; DB 22; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 NVLQ 6  
 DB 1 NVLQ 4  
 RESULT 8  
 AAB72821  
 ID AAB72821 standard; Peptide; 7 AA.  
 XX  
 AC AAB72821;  
 XX  
 DT 09-MAY-2001 (first entry)  
 XX  
 DE Antibacterial peptide SEQ ID NO: 42.  
 XX  
 KW Antibacterial compound; Gram-negative bacterium; pilus; chaperone;  
 KW biofilm; disease treatment; bacterial infection.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200110386-A2.  
 XX  
 PD 15-FEB-2001.  
 XX  
 PF 11-AUG-2000; 2000WO-US22087.  
 XX  
 PR 11-AUG-1999; 99US-0148280.  
 XX  
 PA (UNIW ) UNIV WASHINGTON.  
 XX

PI Hultgren SJ, Sauer FG, Waksman G, Fuetterer K;  
 XX WPI; 2001-226496/23.  
 XX  
 XX An isolated compound for inhibiting pilus assembly -  
 XX  
 XX Claim 12; Page 140; 144pp; English.

XX The present invention provides antibacterial compounds which are able to  
 CC interfere with Gram-negative bacteria pilus formation and assembly, and  
 CC pilus interaction with chaperone proteins. These are useful in the  
 CC treatment of bacterial infection, and in the prevention of biofilm  
 CC formation. They are particularly useful against *Escherichia coli*,  
 CC *Haemophilus influenzae*, *Salmonella enteritidis*, *S. typhimurium*,  
 CC *Bordetella pertussis*, *Yersinia enterocolitica*, *Helicobacter pylori* and  
 CC *Klebsiella pneumoniae*.  
 XX

SQ Sequence 7 AA;

Query Match 52.8%; Score 19; DB 22; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLQ 6  
 DB 1 NVLQ 4

RESULT 9  
 AAB72823  
 ID AAB72823 standard; Peptide; 7 AA.

XX AC AAB72823;  
 XX DT 09-MAY-2001 (first entry)

XX DE Antibacterial peptide SEQ ID NO: 44.

XX KW Antibacterial compound; Gram-negative bacterium; pilus; chaperone;  
 XX biofilm; disease treatment; bacterial infection.

XX OS Synthetic.

XX PN W0200110386-A2.

XX PD 15-FEB-2001.

XX PF 11-AUG-2000; 2000WO-US22087.

XX PR 11-AUG-1999; 99US-0148280.

XX PA (UNIW) UNIV WASHINGTON.

XX PI Hultgren SJ, Sauer FG, Waksman G, Fuetterer K;

XX WPI; 2001-226496/23.

XX An isolated compound for inhibiting pilus assembly -

XX Claim 12; Page 141; 144pp; English.

XX The present invention provides antibacterial compounds which are able to  
 CC interfere with Gram-negative bacteria pilus formation and assembly, and  
 CC pilus interaction with chaperone proteins. These are useful in the  
 CC treatment of bacterial infection, and in the prevention of biofilm  
 CC formation. They are particularly useful against *Escherichia coli*,  
 CC *Haemophilus influenzae*, *Salmonella enteritidis*, *S. typhimurium*,  
 CC *Bordetella pertussis*, *Yersinia enterocolitica*, *Helicobacter pylori* and  
 CC *Klebsiella pneumoniae*.  
 XX

SQ Sequence 7 AA;

Query Match 52.8%; Score 19; DB 22; Length 7;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLQ 6  
 DB 1 NVLQ 4

RESULT 10  
 AAR54171  
 ID AAR54171 standard; Protein; 5 AA.

XX AC AAR54171;

XX DT 22-FEB-1995 (first entry)

XX DE CHA255 light chain CDR2 clone 4.

XX KW Polymerase chain reaction; primer; PCR; amplify; heavy; light;  
 XX chain; complementarity determining region; CDR; variable; constant;  
 XX region; monoclonal antibody; MAB; binding affinity; EDTA; DOTA;  
 XX tumour; cancer; colorectal; breast; metal chelate; hapten.

XX OS Synthetic.

XX PN AU9350602-A.

XX PD 26-MAY-1994.

XX PF 10-NOV-1993; 93AU-0050602.

XX PR 12-NOV-1992; 92US-0975230.

XX PA (HYBR-) HYBRITECH INC.

XX PI Ahrweiler PM, Moore MD;

XX WPI; 1994-209063/26.

XX N-PSDB; AAQ68773.

XX Polypeptide used in imaging and treatment of carcinomas and  
 XX tumours - comprising substd antibody CDR having binding affinity  
 XX for metal chelate of EDTA or DOTA or analogues

XX Claim 25; Fig 3B; 61pp; English.

XX The sequences given in AAR54167-76 represent the wild type and  
 CC mutagenised versions of the complementarity determining region 2 (CDR2)  
 CC of the antibody designated CHA255 light chain. CHA255 is a murine  
 CC monoclonal antibody (MAB) which is capable of binding complexes.  
 CC Mutagenesis of these CDRs, causes the production of polypeptides with a  
 CC particularly high binding affinity for EDTA or DOTA metal complexes.  
 CC CDR1 and -3 of the heavy chain, and CDR2 and -3 of the light chain were  
 CC targeted for mutagenesis. Five residues of both CDR1 and -3 of the  
 CC CHA255 heavy chain, five of seven residues of light chain CDR and six of  
 CC nine light chain CDR3 residues were specifically targeted for  
 CC codon-based mutagenesis. The mutagenised MAB's can be used in  
 CC compositions for in vivo imaging of malignant tissues or tumours. They  
 CC are also useful for the treatment of malignant tissues or tumours eg.  
 CC colorectal or breast cancer. Both methods involve the use of  
 CC radionuclides which bind to metal chelates or haptens which are  
 CC specifically delivered to the target site by a targeting molecule. CDR  
 CC derived peptides may be used to construct bi-functional antibodies  
 CC having dual specificities, or as donor or recipients of CDR sequences.

SQ Sequence 5 AA;

Query Match 50.0%; Score 18; DB 15; Length 5;  
 Best Local Similarity 80.0%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVL 5  
 IIII

Db 1 TNNRL 5

RESULT 11  
AAB52719  
ID AAB52719 standard; Peptide; 5 AA.  
XX AC AAB52719;  
XX DT 23-FEB-2001 (first entry)  
XX DE T21/DP107 peptide fragment #1.  
XX KW Antinflammatory; T21/DP107; gp41 ectodomain; HIV-1 fusion;  
KW formyl peptide receptor family; FPR; inflammatory response up-regulation;  
KW chemoattractant.  
XX OS Human immunodeficiency virus type 1.  
XX PN WO200066622-A1.  
XX PD 09-NOV-2000.  
XX PF 05-MAY-2000; 2000WO-US12371.  
XX PR 05-MAY-1999; 99US-0132686.  
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX PI Wang JM, Oppenheim JJ, Su S, Gong W, Gao J, Murphy PM;  
XX WPI; 2000-656493/63.  
XX DR Administration of peptide agents with a sequence corresponding to a  
PT partial-length T20/DP178 or T21/DP107 formyl peptide receptor  
PT antagonist is used to modulate inflammation -  
XX Claim 12; Page 27; 148pp; English.  
XX The present sequence is a peptide fragment of T21/DP107. T21/DP107 is a  
CC helical segment of the ectodomain of HIV-1 protein gp41. T21/DP107 is  
CC located in the amino terminus of the gp41 ectodomain. gp41 plays a  
CC critical role in the fusion of HIV-1 and host cell membranes. T21/DP107  
CC interacts with members of the formyl peptide receptor (FPR) family and  
CC thereby up-regulates an inflammatory response, and acts as a potent  
CC chemoattractant and activator of human peripheral blood phagocytes  
CC (but not T cells). The present peptide can be used to modulate an  
CC inflammatory response in a subject.  
XX Sequence 5 AA;  
SQ

Query Match 50.0%; Score 18; DB 21; Length 5;  
Best Local Similarity 60.0%; Pred. No. 7.8e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 2 NNVLQ 6  
DB 1 NNLLR 5

RESULT 12  
AAB14673  
ID AAB14673 standard; peptide; 5 AA.  
XX AC AAB14673;  
XX DT 24-NOV-2000 (first entry)  
XX DE HIV-1 isolate LAI gp41 N-helical domain peptide P-17, fragment #2.  
XX KW HIV-1; gp41 N-helical domain; gp41 heptad repeat region;  
KW coiled coil; core 6-helix bundle; viral entry inhibition; immunogenic;  
KW antibody; humoral response; broad spectrum vaccine; anti-HIV;

KW envelope glycoprotein; prophylaxis; therapy; group M; subtype B;  
KW isolate LAI.  
XX Human immunodeficiency virus type 1.  
XX PN WO200040616-A1.  
XX PD 13-JUL-2000.  
XX PF 10-JAN-2000; 2000WO-US00456.  
XX PR 08-JAN-1999; 99US-0115404.  
XX PF 07-JAN-2000; 2000US-0480336.  
XX PA (WILD/) WILD C T.  
XX PA (WEIS/) WEISS C D.  
XX PI Wild CT, Weiss CD;  
XX DR WPI; 2000-465959/40.  
XX Raising neutralizing antibody response to human immunodeficiency virus,  
PT comprises administering a polypeptide capable of forming a stable  
PT coiled-coil solution structure -  
XX Disclosure; Page 35; 97pp; English.  
XX Sequences AAB14672-B14739 represent peptides derived from the  
CC N-helical domain of the gp41 envelope glycoprotein from HIV-1  
CC group M, subtype B, isolate LAI. The invention relates to raising  
CC a neutralising antibody response to a broad spectrum of HIV (human  
CC immunodeficiency virus) strains and isolates, comprising the  
CC administration of a peptide which corresponds to or mimics highly  
CC conserved portions of gp41 which are important in mediating the  
CC process of viral entry into host cells. Such peptides can correspond  
CC to or mimic the coiled coil solution structure of the N-helical domain  
CC (the heptad repeat region), or can correspond or mimic the C-helical  
CC domain (the transmembrane-proximal amphipathic alpha-helical segment),  
CC or the gp41 core 6-helix bundle, which is formed by the interaction  
CC of the N- and C-helical domains of three gp41 proteins. The peptides  
CC can be administered either singly or as a combination (particularly  
CC a combination of N-helical and C-helical peptides), and can be  
CC multimerised. For example, N- and C-helical domain peptides can be  
CC alternately linked together to form a peptide which mimics the core  
CC 6-helix bundle. Administration of the peptide(s) generates a humoral  
CC response, with the production of antibodies against gp41 structures  
CC involved in viral entry. As these portions of gp41 are well conserved,  
CC such antibodies may be effective against a broad range of HIV strains  
CC and isolates. The peptide compositions may be administered as a  
CC prophylactic or therapeutic vaccine to generate antibodies which reduce  
CC or inhibit the ability of HIV to infect uninfected cells. A composition  
CC comprising polyclonal or monoclonal antibodies can be administered to  
CC reduce HIV infection of uninfected cells. Antibodies raised against  
CC entry-relevant gp41 structures may also be used therapeutically and as  
CC tools to further elucidate the mechanism of HIV cell entry.  
XX Sequence 5 AA;  
SQ

Query Match 50.0%; Score 18; DB 21; Length 5;  
Best Local Similarity 60.0%; Pred. No. 7.8e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 2 NNVLQ 6  
DB 1 NNLLR 5

RESULT 13  
AAB54937  
ID AAB54937 standard; Peptide; 5 AA.  
XX AC AAB54937;  
XX XX

DT XX 05-MAR-2001 (first entry)

DE XX Anti-HIV peptide DPl07 carboxy truncation peptide #2.

KW XX Long lasting fusion peptide inhibitor; viral infection; antiviral;

KW XX antifusogenic; mobile blood component; measles virus; MeV; SIV;

KW XX simian immunodeficiency virus; human parainfluenza virus; HPIV; RSV;

KW XX human respiratory syncytial virus; human immunodeficiency virus; HIV.

OS XX Human immunodeficiency virus type 1.

XX XX WO200069902-A1.

PN XX 23-NOV-2000.

XX XX 17-MAY-2000; 2000WO-US13651.

XX XX 17-MAY-1999; 99US-0134406.

PR XX 10-SEP-1999; 99US-0153406.

XX XX (CONJ-) CONJUCHEM INC.

PA XX Bridon DP, Dufresne RP, Boudjellab N, Robitaille M, Milner PG;

XX XX WPI; 2001-007496/01.

DR XX

XX XX A modified peptide and a reactive group which is reactive with amino

PT XX groups, hydroxyl groups, or thiol groups on blood components to form

PT XX stable covalent bonds useful for treatment of viral infections, e.g.

PT XX human immunodeficiency virus

XX XX Disclosure; Page 136; 21pp; English.

PS XX

CC XX The present invention describes a modified anti-viral peptide (I)

CC XX comprising a peptide that exhibits anti-viral activity and a reactive

CC XX group which is reactive with amino groups, hydroxyl groups, or thiol

CC XX groups on blood components to form stable covalent bonds. (I) has

CC XX anti-viral and anti-fusogenic activities. (I) inhibits viral infection

CC XX of cells by inhibiting cell-cell fusion or free virus infection or to

CC XX reduce the level of membrane fusion events between two or more entities,

CC XX e.g., virus-cell or cell-cell, relative to the level of membrane fusion

CC XX that occurs in the absence of the peptide. (I) is useful in the

CC XX treatment of patients who are suffering from viral infection, e.g. HIV,

CC XX RSV, HPIV, MeV and SIV. (I) may be administered prophylactically to

CC XX previously uninfected individuals. This is useful in cases where an

CC XX individual has been subjected to a high risk of exposure to a virus.

CC XX By bonding of long-lived components of the blood, such as immunoglobulin,

CC XX serum albumin, red blood cells and platelets the activity is extended

CC XX for days to weeks. This is due to improved stability in vivo and a

CC XX reduced susceptibility to peptidase or protease degradation. This

CC XX minimises the need for more frequent, or even continual, administration

CC XX of the peptides. AAB54784 to AAB55431 represent peptides used in the

CC XX exemplification of the present invention.

XX XX

SQ XX Sequence 5 AA;

Query Match 50.0%; Score 18; DB 22; Length 5;

Best Local Similarity 60.0%; Pred. No. 7.8e+05;

Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNVLQ 6

DB 1 NNLLR 5

RESULT 14

AAY23409

ID AAY23409 standard; Peptide; 6 AA.

XX AAY23409;

AC AAY23409;

XX

DT 02-SEP-1999 (first entry)

XX

DE XX V beta 6 clone found in MS patients after vaccination with TCR.

XX XX Vaccine; T cell receptor; TCR; T cell; V beta 6.2/3; V beta 6/5;

KW XX V beta 6.7; V beta 2; V beta 5/1; V beta 7; V beta 13; V beta 6;

KW XX multiple sclerosis.

XX XX Synthetic.

OS XX Homo sapiens.

XX XX WO9927957-A1.

PN XX 10-JUN-1999.

XX XX 03-DEC-1997; 97WO-US23147.

XX XX 03-DEC-1997; 97WO-US23147.

PR XX (IMMU-) IMMUNE RESPONSE CORP.

XX XX (KIMM-) KIMMEL CANCER CENT SIDNEY.

PA XX

XX XX Brostoff SW, Carlo DJ, Gold DP, Smith LR, Wilson DB;

XX XX WPI; 1999-404801/34.

DR XX

XX XX T0 cell receptor peptide-derived vaccines

PT XX

PS XX Example 11; Page 85; 104pp; English.

XX XX

CC XX The specification describes vaccines which comprise immunologically

CC XX effective amounts of T cell receptor (TCR) peptides. The TCRs are

CC XX present on the surface of T cells. The TCRs are chosen from V beta

CC XX 6.2/3, V beta 6/5, V beta 6.7, V beta 2, V beta 5/1, V beta 7 or V beta

CC XX 13. The V beta TCR peptide-based vaccines are useful for prevention or

CC XX treatment of multiple sclerosis (MS). The presence of V beta 6.7 appears

CC XX to be particularly associated with multiple sclerosis and can be used

CC XX to determine an individual's susceptibility to multiple sclerosis.

CC XX Vaccinating, rather than passively administering heterologous

CC XX antibodies, allows the host's own immune system to mobilize and suppress

CC XX auto aggressive T cells. Therefore, the suppression is persistent and

CC XX may involve any and all immunological mechanisms in effecting that

CC XX suppression. Such a multi-faceted response is more effective than

CC XX the uni-dimensional suppression achieved by passive administration of

CC XX monoclonal antibodies or extant-derived regulatory T cell clones.

CC XX AAY23387-Y23480 represent peptides derived from TCR V beta 6 clones

CC XX found in the cerebrospinal fluid (CSF) of MS patients, after vaccination

CC XX with V beta 6.

XX XX

SQ XX Sequence 6 AA;

Query Match 50.0%; Score 18; DB 20; Length 6;

Best Local Similarity 60.0%; Pred. No. 7.8e+05;

Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNVLQ 6

DB 1 NNALE 5

RESULT 15

AAW78331

ID AAW78331 standard; peptide; 6 AA.

XX AAW78331;

AC AAW78331;

XX XX

DT 27-APR-1999 (first entry)

XX XX Human fibrinogen gamma A/B chain protein fragment.

DE XX

XX XX Expression; epitope; recombinant vector; structural gene; bacteriophage;

KW XX identification.

XX XX

OS XX Homo sapiens.

XX XX

```

PN US5866363-A.
XX
PD 02-FEB-1999.
XX
PF 28-FEB-1991; 91US-0662764.
XX
PR 28-FEB-1991; 91US-0662764.
PR 28-AUG-1985; 85US-0770390.
PR 26-MAY-1988; 88US-0201358.
XX
PA (PIEC/) PIECZENIK G.
XX
PI Pieczenik G;
XX
DR WPI: 1999-141923/12.
DR N-PSDB; AAX16850.
XX
PT Population of recombinant vectors, oligonucleotides and random
PT peptide sequences - useful for the identification and
PT characterisation of peptide epitopes
XX
PS Example 4; Column 21; 26pp; English.
XX
CC The invention relates to the expression of epitope peptides from
CC recombinant vectors which contain a recombinant structural gene into
CC which a random oligonucleotide of sequence GATCCTTN(15)AA has been
CC inserted. Sequences X16832-X16845 represent examples of the inserted
CC oligonucleotides. Several of the peptides encoded by the oligonucleotides
CC have been shown to have similarity to proteins in the GenBank sequence
CC database. This sequence is a fragment of the human fibrinogen gamma A/B
CC chain precursor protein and is example of a homologous peptide sequence.
CC The recombinant peptides are then expressed on the surface of an organism
CC e.g. a bacteriophage, for ease of identification. The invention may be
CC used for the identification and characterisation of peptide epitopes.
XX
SQ Sequence 6 AA;

Query Match 50.0%; Score 18; DB 20; Length 6;
Best Local Similarity 75.0%; Pred. No. 7.8e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLQ 6
Db 1 NILQ 4

Search completed: February 26, 2003, 15:43:26
Job time : 35 secs

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:45:27 : Search time 12 Seconds  
(without alignments)  
25.148 Million cell updates/sec

Title: US-09-476-485A-24

Perfect score: 36

Sequence: 1 TNNVLQXT 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174566 seqs, 37721826 residues

Total number of hits satisfying chosen parameters: 16185

Minimum DB seq length: 0

Maximum DB seq length: 8

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description       |
|------------|-------|---------------|--------|-------|-------------------|
| 1          | 22    | 61.1          | 6      | 9     | US-09-823-823-61  |
| 2          | 22    | 61.1          | 6      | 10    | US-09-823-829-61  |
| 3          | 20    | 55.6          | 7      | 9     | US-09-823-823-80  |
| 4          | 20    | 55.6          | 7      | 10    | US-09-848-585-22  |
| 5          | 20    | 55.6          | 7      | 10    | US-09-823-829-80  |
| 6          | 17    | 47.2          | 6      | 9     | US-09-727-963A-56 |
| 7          | 17    | 47.2          | 6      | 10    | US-09-990-762-55  |
| 8          | 17    | 47.2          | 7      | 9     | US-09-956-086-18  |
| 9          | 17    | 47.2          | 7      | 9     | US-09-956-087-18  |
| 10         | 17    | 47.2          | 7      | 9     | US-09-995-973-20  |
| 11         | 17    | 47.2          | 7      | 9     | US-09-995-973-24  |
| 12         | 17    | 47.2          | 7      | 9     | US-09-995-973-27  |
| 13         | 17    | 47.2          | 7      | 10    | US-09-865-483-10  |
| 14         | 17    | 47.2          | 8      | 10    | US-09-924-703-12  |
| 15         | 16    | 44.4          | 6      | 10    | US-09-990-762-54  |
| 16         | 16    | 44.4          | 6      | 10    | US-09-911-838-142 |
| 17         | 16    | 44.4          | 6      | 10    | US-09-911-838-144 |
| 18         | 16    | 44.4          | 7      | 10    | US-09-911-838-141 |
| 19         | 16    | 44.4          | 7      | 10    | US-09-911-838-143 |

20 16 44.4 7 10 US-09-911-838-145 Sequence 145, App  
21 16 44.4 8 10 US-09-765-086-157 Sequence 157, App  
22 16 44.4 8 10 US-09-950-313-43 Sequence 43, Appl  
23 15 41.7 4 10 US-09-071-838-231 Sequence 231, App  
24 15 41.7 7 8 US-08-967-573A-22 Sequence 22, Appl  
25 15 41.7 7 9 US-10-080-100-59 Sequence 59, Appl  
26 15 41.7 7 9 US-10-080-100-100 Sequence 100, Appl  
27 15 41.7 7 10 US-09-989-789-3047 Sequence 3047, App  
28 15 41.7 7 10 US-09-989-789-3049 Sequence 3049, App  
29 15 41.7 7 10 US-09-870-379-8 Sequence 8, Appl  
30 15 41.7 8 9 US-10-080-100-45 Sequence 45, Appl  
31 14 38.9 4 10 US-09-220-920-71 Sequence 71, Appl  
32 14 38.9 5 9 US-09-788-006-164 Sequence 164, App  
33 14 38.9 6 10 US-09-947-387-121 Sequence 121, App  
34 14 38.9 6 10 US-09-911-838-140 Sequence 140, App  
35 14 38.9 7 10 US-09-989-789-271 Sequence 271, App  
36 14 38.9 7 10 US-09-989-789-331 Sequence 331, App  
37 14 38.9 7 10 US-09-989-789-370 Sequence 370, App  
38 14 38.9 7 10 US-09-989-789-1013 Sequence 1013, App  
39 14 38.9 7 10 US-09-989-789-1169 Sequence 1169, App  
40 14 38.9 7 10 US-09-900-147-7 Sequence 7, Appl  
41 14 38.9 7 10 US-09-947-387-32 Sequence 32, Appl  
42 14 38.9 7 10 US-09-911-838-139 Sequence 139, App  
43 14 38.9 8 9 US-09-826-290-53 Sequence 53, Appl  
44 14 38.9 8 9 US-09-826-290-135 Sequence 135, App  
45 14 38.9 8 9 US-09-826-290-269 Sequence 269, App

#### ALIGNMENTS

RESULT 1  
US-09-823-823-61  
; Sequence 61, Application US/09823823  
; Patent No. US20020171092A1  
; GENERAL INFORMATION:  
; APPLICANT: Yamamoto, Satoshi  
; APPLICANT: Kasai, Hiroaki  
; APPLICANT: Nakamura, Shoko  
; APPLICANT: Suzuki, Makoto  
; APPLICANT: Hamoda, Tohru  
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION AND DETECTION OF MICROORGANISMS USING  
; FILE REFERENCE: 12817-004001  
; CURRENT FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: US 03/208,688  
; PRIOR FILING DATE: 1998-12-10  
; PRIOR APPLICATION NUMBER: JP 97/343316  
; PRIOR FILING DATE: 1997-12-12  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: Patentin version 2.0  
; SEQ ID NO 61  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetically generated protein  
US-09-823-823-61

Query Match 61.1%; Score 22; DB 9; Length 6;  
Best Local Similarity 66.7%; Pred. No. 1.5e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
|||:|  
Db 1 TNNIPQ 6

RESULT 2  
US-09-823-829-61  
; Sequence 61, Application US/09823829  
; Patent No. US20020146697A1

GENERAL INFORMATION:  
; APPLICANT: Yamamoto, Satoshi  
; APPLICANT: Nakamura, Shoko  
; APPLICANT: Suzuki, Makoto  
; APPLICANT: Kasai, Hiroaki  
; APPLICANT: Hamada, Tohru  
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION AND DETECTION OF MICROORGANISMS  
; FILE REFERENCE: 12817-004001  
; CURRENT APPLICATION NUMBER: US/09/823,829  
; CURRENT FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: US 09/208,688  
; PRIOR FILING DATE: 1998-12-10  
; PRIOR APPLICATION NUMBER: JP 97/343316  
; PRIOR FILING DATE: 1997-12-12  
; NUMBER OF SEQ ID NOS: 82  
; SOFTWARE: PatentIn version 2.0  
; SEQ ID NO 61  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetically generated peptide  
US-09-823-829-61

Query Match 61.1%; Score 22; DB 10; Length 6;  
Best Local Similarity 66.7%; Pred. No. 1.5e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
|||:  
Db 1 TNNIPQ 6

## RESULT 3

US-09-823-823-80  
; Sequence 80, Application US/09823823  
; Patent No. US20020171092A1  
; GENERAL INFORMATION:  
; APPLICANT: Yamamoto, Satoshi  
; APPLICANT: Kasai, Hiroaki  
; APPLICANT: Nakamura, Shoko  
; APPLICANT: Suzuki, Makoto  
; APPLICANT: Hamada, Tohru  
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION AND DETECTION OF MICROORGANISMS USING G  
; FILE REFERENCE: 12817-004001  
; CURRENT APPLICATION NUMBER: US/09/823,823  
; CURRENT FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: US 09/208,688  
; PRIOR FILING DATE: 1998-12-10  
; PRIOR APPLICATION NUMBER: JP 97/343316  
; PRIOR FILING DATE: 1997-12-12  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn version 2.0  
; SEQ ID NO 80  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Xaa at position 6 = Pro or Asn; Xaa at position 7 = Thr or Gln  
US-09-823-823-80

Query Match 55.6%; Score 20; DB 9; Length 7;  
Best Local Similarity 75.0%; Pred. No. 1.5e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4  
|||:  
Db 2 TNNI 5

## RESULT 4

US-09-848-585-22  
; Sequence 22, Application US/09848585  
; Patent No. US20020146391A1  
; GENERAL INFORMATION:  
; APPLICANT: LEVITT, Roy Clifford  
; APPLICANT: MALOY, W. Lee  
; APPLICANT: NICOLAIDES, Nicholas C.  
; TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating Atopic Aller  
; FILE REFERENCE: 36870-5056-12-US  
; CURRENT APPLICATION NUMBER: US/09/848,585  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/002,765  
; PRIOR FILING DATE: 1995-08-24  
; PRIOR APPLICATION NUMBER: US 08/697,419  
; PRIOR FILING DATE: 1996-08-23  
; PRIOR APPLICATION NUMBER: US 08/874,503  
; PRIOR FILING DATE: 1997-06-13  
; PRIOR APPLICATION NUMBER: US 09/325,571  
; PRIOR FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 22  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Residues 8-14 of Mature hIL-9 Receptor  
US-09-848-585-22

Query Match 55.6%; Score 20; DB 10; Length 7;  
Best Local Similarity 75.0%; Pred. No. 1.5e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4  
|||:  
Db 4 TNNI 7

## RESULT 5

US-09-823-829-80  
; Sequence 80, Application US/09823829  
; Patent No. US20020146697A1  
; GENERAL INFORMATION:  
; APPLICANT: Yamamoto, Satoshi  
; APPLICANT: Nakamura, Shoko  
; APPLICANT: Suzuki, Makoto  
; APPLICANT: Kasai, Hiroaki  
; APPLICANT: Hamada, Tohru  
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION AND DETECTION OF MICROORGANISMS  
; FILE REFERENCE: 12817-004001  
; CURRENT APPLICATION NUMBER: US/09/823,829  
; CURRENT FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: US 09/208,688  
; PRIOR FILING DATE: 1998-12-10  
; PRIOR APPLICATION NUMBER: JP 97/343316  
; PRIOR FILING DATE: 1997-12-12  
; NUMBER OF SEQ ID NOS: 82  
; SOFTWARE: PatentIn version 2.0  
; SEQ ID NO 80  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetically generated peptide  
; NAME/KEY: VARIANT  
; LOCATION: 6  
; OTHER INFORMATION: Xaa = Pro or Asn  
; NAME/KEY: VARIANT  
; LOCATION: 7  
; OTHER INFORMATION: Xaa = Thr or Gln

US-09-823-829-80

Query Match 55.6%; Score 20; DB 10; Length 7;  
Best Local Similarity 75.0%; Pred. No. 1.5e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4  
Db 2 TNNI 5

## RESULT 6

US-09-727-963A-56  
; Sequence 56, Application US/09727963A  
; Patent No. US20020155106A1  
; GENERAL INFORMATION:  
; APPLICANT: V.I. Technologies, Inc.  
; APPLICANT: Hammond, David J.  
; TITLE OF INVENTION: METHOD OF IDENTIFYING A LIGAND FOR A TARGET MOLECULE  
; FILE REFERENCE: 18242-505  
; CURRENT APPLICATION NUMBER: US/09/727,963A  
; CURRENT FILING DATE: 2000-12-01  
; PRIOR APPLICATION NUMBER: U.S.N. 60/327,557  
; PRIOR FILING DATE: 1999-12-02  
; NUMBER OF SEQ ID NOS: 99  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 56  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Peptide Ligand  
US-09-727-963A-56

Query Match 47.2%; Score 17; DB 9; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNN 3  
Db 4 TNN 6

## RESULT 7

US-09-990-762-55  
; Sequence 55, Application US/09990762  
; Patent No. US20020119498A1  
; GENERAL INFORMATION:  
; APPLICANT: JOUNG, J. KEITH  
; APPLICANT: MILLER, JEFFREY  
; APPLICANT: PABO, CARL O.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INTERACTION TRAP ASSAYS  
; FILE REFERENCE: MTV-030.02 (20021-3002)  
; CURRENT APPLICATION NUMBER: US/09/990,762  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 09/858,852  
; PRIOR FILING DATE: 2001-05-16  
; PRIOR APPLICATION NUMBER: 60/204,509  
; PRIOR FILING DATE: 2000-05-16  
; NUMBER OF SEQ ID NOS: 97  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 55  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Zinc finger  
; OTHER INFORMATION: recognition sequence  
US-09-990-762-55

Query Match 47.2%; Score 17; DB 10; Length 6;  
Best Local Similarity 75.0%; Pred. No. 1.5e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4  
Db 3 SNNV 6

## RESULT 8

US-09-956-086-18  
; Sequence 18, Application US/09956086  
; Patent No. US20020155498A1  
; GENERAL INFORMATION:  
; APPLICANT: FILPULA, DAVID  
; APPLICANT: WANG, MAOLIANG  
; APPLICANT: SHORR, ROBERT  
; APPLICANT: WHITLOW, MARC  
; APPLICANT: LEE, LIHSYNG S.  
; TITLE OF INVENTION: SINGLE-CHAIN ANTIGEN-BINDING PROTEINS  
; CAPABLE OF GLYCOSYLATION, PRODUCTION AND USES THEREOF  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVE., NW, SUITE 600  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/956,086  
; FILING DATE: 20-Sep-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/069,821  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US 60/063,074  
; FILING DATE: 27-OCT-1997  
; APPLICATION NUMBER: US 60/050,472  
; FILING DATE: 23-JUN-1997  
; APPLICATION NUMBER: US 60/044,449  
; FILING DATE: 30-APR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KIM, JUDITH U.  
; REGISTRATION NUMBER: 40,679  
; REFERENCE/DOCKET NUMBER: 0977.2280003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)371-2600  
; TELEFAX: (202)371-2540  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: No. US20020155498A1 Relevant  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:

US-09-956-086-18 ;  
Query Match 47.2%; Score 17; DB 9; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TNN 3  
Db 3 TNN 5

## RESULT 9

US-09-956-087-18  
; Sequence 18, Application US/09956087

Patent No. US20020161201A1  
GENERAL INFORMATION:  
APPLICANT: FILPULA, DAVID  
WANG, MAOLIANG  
SHORR, ROBERT  
WHITLOW, MARC  
LEE, LIHSYNG S.  
TITLE OF INVENTION: SINGLE-CHAIN ANTIGEN-BINDING PROTEINS  
CAPABLE OF GLYCOSYLATION, PRODUCTION AND USES THEREOF  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., NW, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/956,087  
FILING DATE: 20-Sep-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/069,821  
FILING DATE: 1998-04-30  
APPLICATION NUMBER: US 60/063,074  
FILING DATE: 27-OCT-1997  
APPLICATION NUMBER: US 60/050,472  
FILING DATE: 23-JUN-1997  
APPLICATION NUMBER: US 60/044,449  
FILING DATE: 30-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: KIM, JUDITH U.  
REGISTRATION NUMBER: 40,679  
REFERENCE/DOCKET NUMBER: 0977.2280003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)371-2600  
TELEFAX: (202)371-2540  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-09-956-087-18  
Query Match 47.2%; Score 17; DB 9; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TNN 3  
DB 3 TNN 5  
RESULT 10  
US-09-995-973-20  
Sequence 20, Application US/09995973  
Publication No. US20030024006A1  
GENERAL INFORMATION:  
APPLICANT: CHOO, Yen  
APPLICANT: ULLMAN, Christopher G.  
TITLE OF INVENTION: GENE SWITCHES  
FILE REFERENCE: 8325-2003 / G7-US1  
CURRENT APPLICATION NUMBER: US/09/995,973  
CURRENT FILING DATE: 2002-03-19  
NUMBER OF SEQ ID NOS: 59

SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 20  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: zinc finger  
OTHER INFORMATION: binding domain  
US-09-995-973-20  
Query Match 47.2%; Score 17; DB 9; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TNN 3  
DB 1 TNN 3  
RESULT 11  
US-09-995-973-24  
Sequence 24, Application US/09995973  
Publication No. US20030024006A1  
GENERAL INFORMATION:  
APPLICANT: CHOO, Yen  
APPLICANT: ULLMAN, Christopher G.  
TITLE OF INVENTION: GENE SWITCHES  
FILE REFERENCE: 8325-2003 / G7-US1  
CURRENT APPLICATION NUMBER: US/09/995,973  
CURRENT FILING DATE: 2002-03-19  
NUMBER OF SEQ ID NOS: 59  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 24  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: zinc finger  
OTHER INFORMATION: binding domain  
US-09-995-973-24  
Query Match 47.2%; Score 17; DB 9; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TNN 3  
DB 1 TNN 3  
RESULT 12  
US-09-995-973-27  
Sequence 27, Application US/09995973  
Publication No. US20030024006A1  
GENERAL INFORMATION:  
APPLICANT: CHOO, Yen  
APPLICANT: ULLMAN, Christopher G.  
TITLE OF INVENTION: GENE SWITCHES  
FILE REFERENCE: 8325-2003 / G7-US1  
CURRENT APPLICATION NUMBER: US/09/995,973  
CURRENT FILING DATE: 2002-03-19  
NUMBER OF SEQ ID NOS: 59  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 27  
LENGTH: 7  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: zinc finger  
OTHER INFORMATION: binding domain  
US-09-995-973-27  
Query Match 47.2%; Score 17; DB 9; Length 7;

Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNN 3  
|||  
Db 1 TNN 3

## RESULT 13

US-09-865-483-10  
; Sequence 10, Application US/09865483  
; Patent No. US20020061581A1  
; GENERAL INFORMATION:  
; APPLICANT: LEE, Jong Wook et al.  
; TITLE OF INVENTION: A VARIABLE REGION OF THE MONOCLONAL ANTIBODY AGAINST THE HBV S-S  
; TITLE OF INVENTION: ANTIGEN AND A GENE ENCODING THE SAME  
; FILE REFERENCE: 1599-0197P  
; CURRENT APPLICATION NUMBER: US/09/865,483  
; CURRENT FILING DATE: 2001-03-29  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 10  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-09-865-483-10

Query Match 47.2%; Score 17; DB 10; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNN 3  
|||  
Db 2 TNN 4

## RESULT 14

US-09-924-703-12  
; Sequence 12, Application:US/09924703  
; Patent No. US20020137898A1  
; GENERAL INFORMATION:  
; APPLICANT: TRINCHIERI, GIORGIO  
; APPLICANT: PERUSSIA, BICE  
; APPLICANT: CLARK, STEVEN C.  
; APPLICANT: WONG, GORDON G.  
; APPLICANT: HEWICK, RODNEY  
; APPLICANT: KOBAYASHI, MICHIO  
; TITLE OF INVENTION: ANTIBODIES TO NATURAL KILLER STIMULATORY FACTOR  
; FILE REFERENCE: 01142.0142.0100  
; CURRENT APPLICATION NUMBER: US/09/924,703  
; CURRENT FILING DATE: 2001-08-09  
; PRIOR APPLICATION NUMBER: 09/325,958  
; PRIOR FILING DATE: 1999-06-04  
; PRIOR APPLICATION NUMBER: 08/858,000  
; PRIOR FILING DATE: 1997-05-16  
; PRIOR APPLICATION NUMBER: 08/403,013  
; PRIOR FILING DATE: 1995-03-13  
; PRIOR APPLICATION NUMBER: 07/584,941  
; PRIOR FILING DATE: 1990-09-18  
; PRIOR APPLICATION NUMBER: 07/307,817  
; PRIOR FILING DATE: 1989-02-07  
; PRIOR APPLICATION NUMBER: 07/269,945  
; PRIOR FILING DATE: 1988-11-10  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: peptide  
US-09-924-703-12

Query Match 47.2%; Score 17; DB 10; Length 8;  
Best Local Similarity 60.0%; Pred. No. 1.5e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNVLQ 6  
:|||  
Db 3 SNMLQ 7

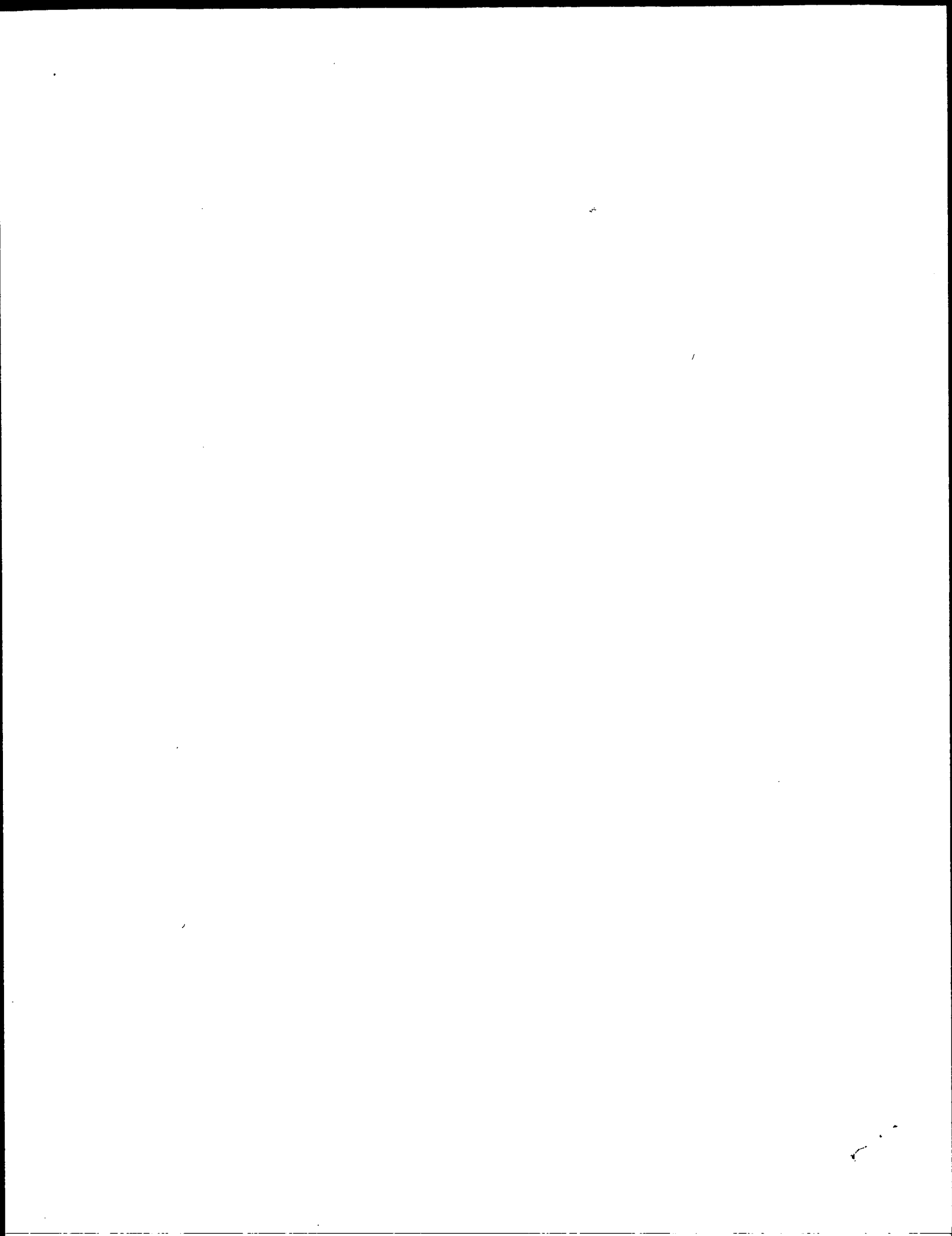
## RESULT 15

US-09-990-762-54  
; Sequence 54, Application US/09990762  
; Patent No. US20020119498A1  
; GENERAL INFORMATION:  
; APPLICANT: JOUNG, J. KEITH  
; APPLICANT: MILLER, JEFFREY  
; APPLICANT: PABO, CARL O.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INTERACTION TRAP ASSAYS  
; FILE REFERENCE: MTV-030.02 (20021-3002)  
; CURRENT APPLICATION NUMBER: US/09/990,762  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 09/858,852  
; PRIOR FILING DATE: 2001-05-16  
; PRIOR APPLICATION NUMBER: 60/204,509  
; PRIOR FILING DATE: 2000-05-16  
; NUMBER OF SEQ ID NOS: 97  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 54  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Zinc finger  
; OTHER INFORMATION: recognition sequence  
US-09-990-762-54

Query Match 44.4%; Score 16; DB 10; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNV 4  
|||  
Db 4 NNV 6

Search completed: February 26, 2003, 15:52:46  
Job time : 13 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:43:32 ; Search time 346 seconds  
(without alignments)  
14.907 Million cell updates/sec

Title: US-09-476-485A-24  
Perfect score: 36  
Sequence: 1 TNNVLOXT 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 223113

Minimum DB seq length: 0  
Maximum DB seq length: 8

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

|     |                                        |
|-----|----------------------------------------|
| 1:  | /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.* |
| 2:  | /cgn2_6/ptodata/1/paa/US06_COMB.pep.*  |
| 3:  | /cgn2_6/ptodata/1/paa/US07_COMB.pep.*  |
| 4:  | /cgn2_6/ptodata/1/paa/US08_COMB.pep.*  |
| 5:  | /cgn2_6/ptodata/1/paa/US081_COMB.pep.* |
| 6:  | /cgn2_6/ptodata/1/paa/US082_COMB.pep.* |
| 7:  | /cgn2_6/ptodata/1/paa/US083_COMB.pep.* |
| 8:  | /cgn2_6/ptodata/1/paa/US084_COMB.pep.* |
| 9:  | /cgn2_6/ptodata/1/paa/US085_COMB.pep.* |
| 10: | /cgn2_6/ptodata/1/paa/US086_COMB.pep.* |
| 11: | /cgn2_6/ptodata/1/paa/US087_COMB.pep.* |
| 12: | /cgn2_6/ptodata/1/paa/US088_COMB.pep.* |
| 13: | /cgn2_6/ptodata/1/paa/US089_COMB.pep.* |
| 14: | /cgn2_6/ptodata/1/paa/US090_COMB.pep.* |
| 15: | /cgn2_6/ptodata/1/paa/US091_COMB.pep.* |
| 16: | /cgn2_6/ptodata/1/paa/US092_COMB.pep.* |
| 17: | /cgn2_6/ptodata/1/paa/US093_COMB.pep.* |
| 18: | /cgn2_6/ptodata/1/paa/US094_COMB.pep.* |
| 19: | /cgn2_6/ptodata/1/paa/US095_COMB.pep.* |
| 20: | /cgn2_6/ptodata/1/paa/US096_COMB.pep.* |
| 21: | /cgn2_6/ptodata/1/paa/US097_COMB.pep.* |
| 22: | /cgn2_6/ptodata/1/paa/US098_COMB.pep.* |
| 23: | /cgn2_6/ptodata/1/paa/US099_COMB.pep.* |
| 24: | /cgn2_6/ptodata/1/paa/US100_COMB.pep.* |
| 25: | /cgn2_6/ptodata/1/paa/US101_COMB.pep.* |
| 26: | /cgn2_6/ptodata/1/paa/US102_COMB.pep.* |
| 27: | /cgn2_6/ptodata/1/paa/US60_COMB.pep.*  |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES  |       |                    |    |             |                     |
|------------|-------|--------------------|----|-------------|---------------------|
| Result No. | Score | Query Match Length | ID | Description |                     |
| 1          | 34    | 94.4               | 8  | 1           | PCT-US98-13046-24   |
| 2          | 34    | 94.4               | 8  | 18          | US-09-476-485A-24   |
| 3          | 34    | 94.4               | 8  | 24          | US-10-045-353-24    |
| 4          | 25    | 69.4               | 8  | 1           | PCT-US01-08656-8073 |
| 5          | 22    | 61.1               | 6  | 22          | US-09-823-823-61    |
| 6          | 22    | 61.1               | 6  | 22          | US-09-823-823-61    |

|    |      |   |    |                     |                    |
|----|------|---|----|---------------------|--------------------|
| 7  | 55.6 | 7 | 10 | US-08-697-360-22    | Sequence 22, Appl  |
| 8  | 55.6 | 7 | 10 | US-08-697-419-22    | Sequence 22, Appl  |
| 9  | 55.6 | 7 | 10 | US-08-697-440-22    | Sequence 22, Appl  |
| 10 | 55.6 | 7 | 10 | US-08-697-471-22    | Sequence 22, Appl  |
| 11 | 55.6 | 7 | 10 | US-08-697-471B-22   | Sequence 22, Appl  |
| 12 | 55.6 | 7 | 10 | US-08-697-472-22    | Sequence 22, Appl  |
| 13 | 55.6 | 7 | 10 | US-08-697-473-22    | Sequence 22, Appl  |
| 14 | 55.6 | 7 | 11 | US-08-702-105-22    | Sequence 22, Appl  |
| 15 | 55.6 | 7 | 11 | US-08-702-110-22    | Sequence 22, Appl  |
| 16 | 55.6 | 7 | 11 | US-08-702-168-22    | Sequence 22, Appl  |
| 17 | 55.6 | 7 | 12 | US-08-874-503-22    | Sequence 22, Appl  |
| 18 | 55.6 | 7 | 22 | US-09-823-823-80    | Sequence 80, Appl  |
| 19 | 55.6 | 7 | 22 | US-09-823-829-80    | Sequence 80, Appl  |
| 20 | 55.6 | 7 | 22 | US-09-848-585-22    | Sequence 22, Appl  |
| 21 | 55.6 | 8 | 16 | US-09-206-786-59    | Sequence 59, Appl  |
| 22 | 55.6 | 8 | 16 | US-09-206-786A-59   | Sequence 59, Appl  |
| 23 | 55.6 | 8 | 18 | US-09-458-297-113   | Sequence 113, Appl |
| 24 | 55.6 | 8 | 18 | US-09-458-297-306   | Sequence 306, Appl |
| 25 | 55.6 | 8 | 18 | US-09-458-297A-113  | Sequence 113, Appl |
| 26 | 55.6 | 8 | 18 | US-09-458-297A-306  | Sequence 306, Appl |
| 27 | 52.8 | 7 | 1  | PCT-US00-06588-36   | Sequence 36, Appl  |
| 28 | 52.8 | 7 | 1  | PCT-US00-06588-37   | Sequence 37, Appl  |
| 29 | 52.8 | 7 | 20 | US-09-637-216-1     | Sequence 1, Appl   |
| 30 | 52.8 | 7 | 20 | US-09-637-216-42    | Sequence 42, Appl  |
| 31 | 52.8 | 7 | 20 | US-09-637-216-44    | Sequence 44, Appl  |
| 32 | 52.8 | 8 | 18 | US-09-412-863-1681  | Sequence 1681, Ap  |
| 33 | 52.8 | 8 | 18 | US-09-412-863-4370  | Sequence 4370, Ap  |
| 34 | 52.8 | 8 | 18 | US-09-412-863-6754  | Sequence 6754, Ap  |
| 35 | 50.0 | 5 | 8  | US-08-406-779-89    | Sequence 89, Appl  |
| 36 | 50.0 | 5 | 19 | US-09-515-965A-1690 | Sequence 1690, Ap  |
| 37 | 50.0 | 5 | 23 | US-09-350-369C-547  | Sequence 547, Appl |
| 38 | 50.0 | 5 | 24 | US-10-005-305-98    | Sequence 98, Appl  |
| 39 | 50.0 | 6 | 13 | US-08-984-098-106   | Sequence 106, Appl |
| 40 | 50.0 | 6 | 19 | US-09-515-965A-1691 | Sequence 1691, Ap  |
| 41 | 50.0 | 6 | 23 | US-09-350-369C-548  | Sequence 548, Appl |
| 42 | 50.0 | 6 | 24 | US-10-005-305-99    | Sequence 99, Appl  |
| 43 | 50.0 | 7 | 8  | US-08-407-620A-5    | Sequence 5, Appl   |
| 44 | 50.0 | 7 | 19 | US-09-515-965A-1692 | Sequence 1692, Ap  |
| 45 | 50.0 | 7 | 20 | US-09-637-216-33    | Sequence 33, Appl  |

## ALIGNMENTS

RESULT 1  
; Sequence 24, Application PC/TUS9813046  
; GENERAL INFORMATION:  
; APPLICANT: Colucci et al.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hofmann & Baron, LLP  
; STREET: 6900 Jericho Turnpike  
; CITY: Syosset  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11791  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US98/13046  
; FILING DATE: June 23, 1998  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/881,189  
; FILING DATE: June 24, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601

```
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
PCT-US98-13046-24

Query Match          94.4%; Score 34; DB 1; Length 8;
Best Local Similarity 87.5%; Pred. No. 4.2e+06;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
   |||||
Db 1 TNNVLQVT 8

RESULT 2
US-09-476-485A-24
; Sequence 24, Application US/09476485A
; GENERAL INFORMATION:
; APPLICANT: Colucci, M. Gabriella
; APPLICANT: Chrispeels, Maarten J.
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for
; TITLE OF INVENTION: and Products of Their Use
; FILE REFERENCE: 108236.119
; CURRENT APPLICATION NUMBER: US/09/476,485A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/881,189
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 24
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Dolichos lablab
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (7)..(7)
; OTHER INFORMATION: Amino acid 7 is Xaa wherein Xaa = any amino acid.
US-09-476-485A-24

Query Match          94.4%; Score 34; DB 18; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
   |||||
Db 1 TNNVLQXT 8

RESULT 3
US-10-045-353-24
; Sequence 24, Application US/10045353
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; PROGENITOR CELL PRESERVATION FACTOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
```

```
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/045,353
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/881,189
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-10-045-353-24

Query Match          94.4%; Score 34; DB 24; Length 8;
Best Local Similarity 87.5%; Pred. No. 4.2e+06;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
   |||||
Db 1 TNNVLQVT 8

RESULT 4
PCT-US01-08656-8073
; Sequence 8073, Application PC/TUS0108656
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: PCT/US01/08656
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 8073
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-08656-8073

Query Match          69.4%; Score 25; DB 1; Length 8;
Best Local Similarity 71.4%; Pred. No. 4.2e+06;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NNVLQXT 8
   |||
Db 1 NNVLKFT 7

RESULT 5
US-09-823-823-61
; Sequence 61, Application US/09823823
; GENERAL INFORMATION:
; APPLICANT: Yamamoto, Satoshi
; APPLICANT: Kasai, Hiroaki
; APPLICANT: Nakamura, Shoko
; APPLICANT: Suzuki, Makoto
; APPLICANT: Hamada, Tohru
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION AND DETECTION OF MICROORGANISMS USING
```



; TITLE OF INVENTION: GENE AS AN INDICATOR  
 ; FILE REFERENCE: 12817-004001  
 ; CURRENT APPLICATION NUMBER: US/09/823,823  
 ; CURRENT FILING DATE: 2001-03-30  
 ; PRIOR APPLICATION NUMBER: US 09/208,688  
 ; PRIOR FILING DATE: 1998-12-10  
 ; PRIOR APPLICATION NUMBER: JP 97/343316  
 ; PRIOR FILING DATE: 1997-12-12  
 ; NUMBER OF SEQ ID NOS: 80  
 ; SOFTWARE: PatentIn version 2.0  
 ; SEQ ID NO 61  
 ; LENGTH: 6  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Synthetically generated protein  
 US-09-823-823-61

Query Match 61.1%; Score 22; DB 22; Length 6;  
 Best Local Similarity 66.7%; Pred. No. 4.2e+06;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
 Db 1 TNNIPQ 6

## RESULT 6

US-09-823-829-61  
 ; Sequence 61, Application US/09823829  
 ; GENERAL INFORMATION:

; APPLICANT: Yamamoto, Satoshi  
 ; APPLICANT: Nakamura, Shoko  
 ; APPLICANT: Suzuki, Makoto  
 ; APPLICANT: Kasai, Hiroaki  
 ; APPLICANT: Hamada, Tohru

; TITLE OF INVENTION: METHOD FOR IDENTIFICATION AND DETECTION OF MICROORGANISMS  
 ; FILE REFERENCE: 12817-004001

; CURRENT APPLICATION NUMBER: US/09/823,829  
 ; CURRENT FILING DATE: 2001-03-30

; PRIOR APPLICATION NUMBER: US 09/208,688  
 ; PRIOR FILING DATE: 1998-12-10

; PRIOR APPLICATION NUMBER: JP 97/343316  
 ; PRIOR FILING DATE: 1997-12-12

; NUMBER OF SEQ ID NOS: 82  
 ; SOFTWARE: PatentIn version 2.0

; SEQ ID NO 61  
 ; LENGTH: 6

; TYPE: PRT  
 ; ORGANISM: Artificial Sequence

; FEATURE:  
 ; OTHER INFORMATION: synthetically generated peptide

US-09-823-829-61

Query Match 61.1%; Score 22; DB 22; Length 6;  
 Best Local Similarity 66.7%; Pred. No. 4.2e+06;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
 Db 1 TNNIPQ 6

## RESULT 7

US-08-697-360-22

; Sequence 22, Application US/08697360  
 ; GENERAL INFORMATION:

; APPLICANT: Levitt, Roy C.  
 ; APPLICANT: Maloy, W. Lee  
 ; APPLICANT: Kari, Prasad

; TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating

; TITLE OF INVENTION: Atopic Allergies Including Asthma And Related Disorders.  
 ; NUMBER OF SEQUENCES: 22  
 ; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner  
 ; STREET: 1300 I Street, N.W., Suite 700  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: USA

; ZIP: 20005-3315  
 ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/697,360  
 ; FILING DATE: 23-AUG-1996

; CLASSIFICATION: 424  
 ; ATTORNEY/AGENT INFORMATION:

; NAME: Fordis, Jean B.  
 ; REGISTRATION NUMBER: 32,984

; REFERENCE/DOCKET NUMBER: 05387.0056-02000  
 ; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)408-4000  
 ; TELEFAX: (202)408-4400

; INFORMATION FOR SEQ ID NO: 22:  
 ; SEQUENCE CHARACTERISTICS:

; LENGTH: 7 amino acids  
 ; TYPE: amino acid

; STRANDEDNESS: single  
 ; TOPOLOGY: linear

; MOLECULE TYPE: peptide  
 US-08-697-360-22

## Query Match

55.6%; Score 20; DB 10; Length 7;  
 Best Local Similarity 75.0%; Pred. No. 4.2e+06;

Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4  
 Db 4 TNNI 7

## RESULT 8

US-08-697-419-22

; Sequence 22, Application US/08697419  
 ; GENERAL INFORMATION:

; APPLICANT: Levitt, Roy C.  
 ; APPLICANT: Maloy, W. Lee

; APPLICANT: Kari, Prasad  
 ; APPLICANT: Nicolaides, Nicholas C.

; TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating  
 ; TITLE OF INVENTION: Atopic Allergies Including Asthma And Related Disorders.

; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner  
 ; STREET: 1300 I Street, N.W., Suite 700

; CITY: Washington  
 ; STATE: D.C.

; COUNTRY: USA  
 ; ZIP: 20005-3315

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/697,419  
 ; FILING DATE: 23-AUG-1996

; CLASSIFICATION: 536  
 ; ATTORNEY/AGENT INFORMATION:

; NAME: Fordis, Jean B.  
 ; REGISTRATION NUMBER: 32,984

REFERENCE/DOCKET NUMBER: 05387.0056-01000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)408-4000  
TELEFAX: (202)408-4400  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-697-419-22

Query Match 55.6%; Score 20; DB 10; Length 7;  
Best Local Similarity 75.0%; Pred. No. 4.2e+06;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNV 4  
Db 4 TNNI 7

RESULT 9  
US-08-697-440-22  
Sequence 22, Application US/08697440  
GENERAL INFORMATION:  
APPLICANT: Levitt, Roy C.  
APPLICANT: Maloy, W. Lee  
APPLICANT: Kari, Prasad  
APPLICANT: Nicolaides, Nicholas C.  
TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating  
TITLE OF INVENTION: Atopic Allergies Including Asthma And Related Disorders.  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/697,440  
FILING DATE: 23-AUG-1996  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fordis, Jean B.  
REGISTRATION NUMBER: 32,984  
REFERENCE/DOCKET NUMBER: 05387.0056-09000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)408-4000  
TELEFAX: (202)408-4400  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-697-440-22

Query Match 55.6%; Score 20; DB 10; Length 7;  
Best Local Similarity 75.0%; Pred. No. 4.2e+06;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNV 4  
Db 4 TNNI 7

RESULT 10  
US-08-697-471-22  
Sequence 22, Application US/08697471  
GENERAL INFORMATION:  
APPLICANT: Levitt, Roy C.  
APPLICANT: Maloy, W. Lee  
APPLICANT: Kari, Prasad  
APPLICANT: Nicolaides, Nicholas C.  
TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating  
TITLE OF INVENTION: Atopic Allergies Including Asthma And Related Disorders.  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/697,471  
FILING DATE: 23-AUG-1996  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Fordis, Jean B.  
REGISTRATION NUMBER: 32,984  
REFERENCE/DOCKET NUMBER: 05387.0056-05000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)408-4000  
TELEFAX: (202)408-4400  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-697-471-22

Query Match 55.6%; Score 20; DB 10; Length 7;  
Best Local Similarity 75.0%; Pred. No. 4.2e+06;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNV 4  
Db 4 TNNI 7

RESULT 11  
US-08-697-471B-22  
Sequence 22, Application US/08697471B  
GENERAL INFORMATION:  
APPLICANT: Levitt, Roy C.  
APPLICANT: Maloy, W. Lee  
APPLICANT: Kari, U. Prasad  
APPLICANT: Nicolaides, Nicholas C.  
TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating  
TITLE OF INVENTION: Atopic Allergies Including Asthma And Related Disorders.  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

Qy 1 TNNV 4  
Db 4 TNNI 7

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/697,471B  
FILING DATE: 23-AUG-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Fordis, Jean B.  
REGISTRATION NUMBER: 32,984  
REFERENCE/DOCKET NUMBER: 05387.0056-05000  
TELEPHONE: (202)408-4000  
TELEFAX: (202)408-4400  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-697-471B-22

Query Match 55.6%; Score 20; DB 10; Length 7;  
Best Local Similarity 75.0%; Pred. No. 4.2e+06;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4  
Db 4 TNNI 7

RESULT 12  
US-08-697-472-22  
Sequence 22, Application US/08697472  
GENERAL INFORMATION:  
APPLICANT: Levitt, Roy C.  
APPLICANT: Maloy, W. Lee  
APPLICANT: Kari, Prasad  
TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating  
TITLE OF INVENTION: Atopic Allergies Including Asthma And Related Disorders.  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/697,472  
FILING DATE: 23-AUG-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Fordis, Jean B.  
REGISTRATION NUMBER: 32,984  
REFERENCE/DOCKET NUMBER: 05387.0056-04000  
TELEPHONE: (202)408-4000  
TELEFAX: (202)408-4400  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

US-08-697-472-22

Query Match 55.6%; Score 20; DB 10; Length 7;  
Best Local Similarity 75.0%; Pred. No. 4.2e+06;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4  
Db 4 TNNI 7

RESULT 13  
US-08-697-473-22  
Sequence 22, Application US/08697473  
GENERAL INFORMATION:  
APPLICANT: Levitt, Roy C.  
APPLICANT: Maloy, W. Lee  
APPLICANT: Kari, Prasad  
TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating  
TITLE OF INVENTION: Atopic Allergies Including Asthma And Related Disorders.  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/697,473  
FILING DATE: 23-AUG-1996  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Fordis, Jean B.  
REGISTRATION NUMBER: 32,984  
REFERENCE/DOCKET NUMBER: 05387.0056-03000  
TELEPHONE: (202)408-4000  
TELEFAX: (202)408-4400  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-697-473-22

Query Match 55.6%; Score 20; DB 10; Length 7;  
Best Local Similarity 75.0%; Pred. No. 4.2e+06;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4  
Db 4 TNNI 7

RESULT 14  
US-08-702-105-22  
Sequence 22, Application US/08702105  
GENERAL INFORMATION:  
APPLICANT: Levitt, Roy C.  
APPLICANT: Maloy, W. Lee  
APPLICANT: Kari, Prasad  
TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating  
TITLE OF INVENTION: Atopic Allergies Including Asthma And Related Disorders.

```
?
? NUMBER OF SEQUENCES: 22
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
? STREET: 1300 I Street, N.W., Suite 700
? CITY: Washington
? STATE: D.C.
? COUNTRY: USA
? ZIP: 20005-3315
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/702,105
? FILING DATE: 23-AUG-1996
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Fordis, Jean B.
? REGISTRATION NUMBER: 32,984
? REFERENCE/DOCKET NUMBER: 05387.0056-06000
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202)408-4000
? TELEFAX: (202)408-4400
? INFORMATION FOR SEQ ID NO: 22:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 7 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? US-08-702-105-22

Query Match 55.6%; Score 20; DB 11; Length 7;
Best Local Similarity 75.0%; Pred. No. 4.2e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNV 4
Db 4 TNNI 7

RESULT 15
US-08-702-110-22
? Sequence 22, Application US/08702110
? GENERAL INFORMATION:
? APPLICANT: Levitt, Roy C.
? APPLICANT: Maloy, W. Lee
? APPLICANT: Kari, Prasad
? APPLICANT: Nicolaides, Nicholas C.
? TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating
? TITLE OF INVENTION: Atopic Allergies Including Asthma And Related Disorders.
? NUMBER OF SEQUENCES: 22
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
? STREET: 1300 I Street, N.W., Suite 700
? CITY: Washington
? STATE: D.C.
? COUNTRY: USA
? ZIP: 20005-3315
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/702,110
? FILING DATE: 23-AUG-1996
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Fordis, Jean B.
? REGISTRATION NUMBER: 32,984
? REFERENCE/DOCKET NUMBER: 05387.0056-07000
```

```
?
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202)408-4000
? TELEFAX: (202)408-4400
? INFORMATION FOR SEQ ID NO: 22:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 7 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? US-08-702-110-22

Query Match 55.6%; Score 20; DB 11; Length 7;
Best Local Similarity 75.0%; Pred. No. 4.2e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNV 4
Db 4 TNNI 7

Search completed: February 26, 2003, 15:51:57
Job time : 347 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:43:47 ; Search time 23 seconds  
(without alignments)  
31.790 Million cell updates/sec

Title: US-09-476-485A-24  
Perfect score: 36  
Sequence: 1 TNNVLOXT 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 486122 seqs, 91396495 residues

Total number of hits satisfying chosen parameters: 40840

Minimum DB seq length: 0  
Maximum DB seq length: 8

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description          |
|------------|-------|-------------|--------|----|----------------------|
| 1          | 34    | 94.4        | 8      | 6  | US-10-190-258A-11    |
| 2          | 23    | 63.9        | 8      | 5  | US-09-641-528A-4816  |
| 3          | 23    | 63.9        | 8      | 5  | US-09-641-528A-8925  |
| 4          | 23    | 63.9        | 8      | 5  | US-09-641-528A-13537 |
| 5          | 23    | 63.9        | 8      | 5  | US-09-641-528A-16461 |
| 6          | 23    | 63.9        | 8      | 5  | US-09-641-528A-18438 |
| 7          | 23    | 63.9        | 8      | 5  | US-09-641-528A-19958 |
| 8          | 20    | 55.6        | 8      | 5  | US-09-641-528A-21718 |
| 9          | 20    | 55.6        | 8      | 5  | US-09-641-528A-30040 |
| 10         | 19    | 52.8        | 7      | 5  | US-09-637-2168-1     |
| 11         | 19    | 52.8        | 7      | 5  | US-09-637-2168-42    |
| 12         | 19    | 52.8        | 7      | 5  | US-09-637-2168-44    |
| 13         | 18    | 50.0        | 7      | 1  | PCT-US02-35009-11    |
| 14         | 18    | 50.0        | 7      | 5  | US-09-637-2168-33    |
| 15         | 18    | 50.0        | 7      | 5  | US-10-284-252-11     |
| 16         | 18    | 50.0        | 8      | 5  | US-09-641-528A-4696  |
| 17         | 18    | 50.0        | 8      | 5  | US-09-641-528A-8925  |
| 18         | 18    | 50.0        | 8      | 5  | US-09-641-528A-13537 |
| 19         | 18    | 50.0        | 8      | 5  | US-09-641-528A-16461 |
| 20         | 18    | 50.0        | 8      | 5  | US-09-641-528A-18438 |
| 21         | 18    | 50.0        | 8      | 5  | US-09-641-528A-19958 |
| 22         | 18    | 50.0        | 8      | 5  | US-09-641-528A-21718 |
| 23         | 18    | 50.0        | 8      | 5  | US-09-641-528A-30040 |
| 24         | 18    | 50.0        | 8      | 5  | US-09-641-528A-41860 |
| 25         | 18    | 50.0        | 8      | 5  | US-09-641-528A-4696  |
| 26         | 18    | 50.0        | 8      | 5  | US-09-641-528A-4816  |

27 18 50.0 8 5 US-09-641-528A-4816 Sequence 4816, Ap  
28 18 50.0 8 5 US-09-641-528A-8925 Sequence 8925, Ap  
29 18 50.0 8 5 US-09-641-528A-13537 Sequence 13537, A  
30 18 50.0 8 5 US-09-641-528A-16461 Sequence 16461, A  
31 18 50.0 8 5 US-09-641-528A-18438 Sequence 18438, A  
32 18 50.0 8 5 US-09-641-528A-19958 Sequence 19958, A  
33 18 50.0 8 5 US-09-641-528A-21718 Sequence 21718, A  
34 18 50.0 8 5 US-09-641-528A-30040 Sequence 30040, A  
35 18 50.0 8 5 US-09-641-528A-41860 Sequence 41860, A  
36 17 47.2 5 5 US-09-637-2168-1 Sequence 36, Appl  
37 17 47.2 5 5 US-09-637-2168-42 Sequence 37, Appl  
38 17 47.2 5 5 US-09-637-2168-44 Sequence 38, Appl  
39 17 47.2 5 5 US-09-637-2168-46 Sequence 39, Appl  
40 17 47.2 5 5 US-09-637-2168-48 Sequence 40, Appl  
41 17 47.2 5 5 US-09-637-2168-50 Sequence 41, Appl  
42 17 47.2 5 5 US-09-637-2168-52 Sequence 42, Appl  
43 17 47.2 5 5 US-09-637-2168-54 Sequence 43, Appl  
44 17 47.2 5 5 US-09-637-2168-56 Sequence 44, Appl  
45 17 47.2 5 5 US-09-637-2168-58 Sequence 45, Appl

ALIGNMENTS

RESULT 1  
US-10-190-258A-11  
; Sequence 11, Application US/10190258A  
; GENERAL INFORMATION:  
; APPLICANT: Moore, Jeffrey G  
; TITLE OF INVENTION: Dendritic Cell Isolation Methods  
; FILE REFERENCE: 108236.132  
; CURRENT APPLICATION NUMBER: US/10/190,258A  
; CURRENT FILING DATE: 2002-07-03  
; PRIOR APPLICATION NUMBER: US 60/303,265  
; PRIOR FILING DATE: 2001-07-05  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Dolichos lablab  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 7  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-10-190-258A-11

Query Match 94.4%; Score 34; DB 6; Length 8;  
Best Local Similarity 100.0%; Pred. No. 4e+05; 0;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TNNVLOXT 8  
Db 1 TNNVLOXT 8  
RESULT 2  
US-09-641-528A-4836  
; Sequence 4836, Application US/09641528  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641,528  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10

; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 4836  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528-4836

Query Match 63.9%; Score 23; DB 5; Length 8;  
Best Local Similarity 66.7%; Pred. No. 4e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
| | | |  
Db 2 TRNLIQ 7

## RESULT 3

US-09-641-528-18778  
; Sequence 18778, Application US/09641528  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641,528  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 18778  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528-18778

Query Match 63.9%; Score 23; DB 5; Length 8;  
Best Local Similarity 66.7%; Pred. No. 4e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
| | | |  
Db 1 TRNLIQ 6

## RESULT 4

US-09-641-528-21808  
; Sequence 21808, Application US/09641528  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641,528  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504

; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 21808  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528-21808

Query Match 63.9%; Score 23; DB 5; Length 8;  
Best Local Similarity 66.7%; Pred. No. 4e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
| | | |  
Db 2 TRNLIQ 7

## RESULT 5

US-09-641-528A-4836  
; Sequence 4836, Application US/09641528A  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641,528A  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 4836  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528A-4836

Query Match 63.9%; Score 23; DB 5; Length 8;  
Best Local Similarity 66.7%; Pred. No. 4e+05;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQ 6  
| | | |  
Db 2 TRNLIQ 7

## RESULT 6

US-09-641-528A-18778  
; Sequence 18778, Application US/09641528A  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641,528A  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 18778  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528A-18778

Query Match 63.9%; Score 23; DB 5; Length 8;  
Best Local Similarity 66.7%; Pred. No. 4e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 1; Gaps 0;

QY 1 TNNVLQ 6  
| | : | |  
Db 1 TRNIIQ 6

RESULT 7  
US-09-641-528A-21808  
; Sequence 21808, Application US/09641528A  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; TITLE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641.528A  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 21808  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528A-21808

Query Match 63.9%; Score 23; DB 5; Length 8;  
Best Local Similarity 66.7%; Pred. No. 4e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 1; Gaps 0;

QY 1 TNNVLQ 6  
| | : | |  
Db 2 TRNIIQ 7

RESULT 8  
US-09-641-528-44907  
; Sequence 44907, Application US/09641528  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; TITLE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641.528  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 44907

; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528-44907

Query Match 55.6%; Score 20; DB 5; Length 8;  
Best Local Similarity 60.0%; Pred. No. 4e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNVILQ 6  
| | | : |  
Db 1 NNVIE 5

RESULT 9  
US-09-641-528A-44907  
; Sequence 44907, Application US/09641528A  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; TITLE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641.528A  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 44907  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528A-44907

Query Match 55.6%; Score 20; DB 5; Length 8;  
Best Local Similarity 60.0%; Pred. No. 4e+05;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNVILQ 6  
| | | : |  
Db 1 NNVIE 5

RESULT 10  
US-09-637-216B-1  
; Sequence 1, Application US/09637216B  
; GENERAL INFORMATION:  
; APPLICANT: WASHINGTON UNIVERSITY  
; TITLE OF INVENTION: ANTI-BACTERIAL COMPOUNDS DIRECTED AGAINST PILUS  
; TITLE OF INVENTION: BIOGENESIS, ADHESION AND ACTIVITY; CO-CRYSTALS OF PILUS  
; TITLE OF INVENTION: SUBUNITS AND METHODS OF USE THEREOF  
; FILE REFERENCE: WSHU2005.1  
; CURRENT APPLICATION NUMBER: US/09/637,216B  
; CURRENT FILING DATE: 2000-08-11  
; PRIOR APPLICATION NUMBER: US 60/148,280  
; PRIOR FILING DATE: 1999-08-11  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthesized

; OTHER INFORMATION: Sequence  
US-09-637-216B-1

Query Match 52.8%; Score 19; DB 5; Length 7;  
Best Local Similarity 100.0%; Pred. No. 4e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLQ 6  
| | | |  
Db 1 NVLQ 4

RESULT 11

US-09-637-216B-42  
; Sequence 42, Application US/09637216B  
; GENERAL INFORMATION:  
; APPLICANT: WASHINGTON UNIVERSITY  
; TITLE OF INVENTION: ANTI-BACTERIAL COMPOUNDS DIRECTED AGAINST PILUS  
; TITLE OF INVENTION: BIOGENESIS, ADHESION AND ACTIVITY; CO-CRYSTALS OF PILUS  
; TITLE OF INVENTION: SUBUNITS AND METHODS OF USE THEREOF  
; FILE REFERENCE: WSHU2005.1  
; CURRENT APPLICATION NUMBER: US/09/637,216B  
; CURRENT FILING DATE: 2000-08-11  
; PRIOR APPLICATION NUMBER: US 60/148,280  
; PRIOR FILING DATE: 1999-08-11  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 42  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthesized  
; OTHER INFORMATION: Sequence  
US-09-637-216B-42

Query Match 52.8%; Score 19; DB 5; Length 7;  
Best Local Similarity 100.0%; Pred. No. 4e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLQ 6  
| | | |  
Db 1 NVLQ 4

RESULT 12

US-09-637-216B-44  
; Sequence 44, Application US/09637216B  
; GENERAL INFORMATION:  
; APPLICANT: WASHINGTON UNIVERSITY  
; TITLE OF INVENTION: ANTI-BACTERIAL COMPOUNDS DIRECTED AGAINST PILUS  
; TITLE OF INVENTION: BIOGENESIS, ADHESION AND ACTIVITY; CO-CRYSTALS OF PILUS  
; TITLE OF INVENTION: SUBUNITS AND METHODS OF USE THEREOF  
; FILE REFERENCE: WSHU2005.1  
; CURRENT APPLICATION NUMBER: US/09/637,216B  
; CURRENT FILING DATE: 2000-08-11  
; PRIOR APPLICATION NUMBER: US 60/148,280  
; PRIOR FILING DATE: 1999-08-11  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 44  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthesized  
; OTHER INFORMATION: Sequence  
US-09-637-216B-44

Query Match 52.8%; Score 19; DB 5; Length 7;  
Best Local Similarity 100.0%; Pred. No. 4e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLQ 6  
| | | |  
Db 1 NVLQ 4

QY 3 NVLQ 6  
| | | |  
Db 1 NVLQ 4

RESULT 13

PCT-US02-35009-11  
; Sequence 11, Application PC/TUS0235009  
; GENERAL INFORMATION:  
; APPLICANT: Jones, C. Hal  
; APPLICANT: Dexter, Paul L.  
; APPLICANT: Evans, Amy K.  
; APPLICANT: Hruby, Dennis E.  
; TITLE OF INVENTION: DEGP Protease: Cleavage Site  
; TITLE OF INVENTION: Identification and Proteolysis of a Natural Target in E.  
; TITLE OF INVENTION: Coli  
; FILE REFERENCE: 016921-169  
; CURRENT APPLICATION NUMBER: PCT/US02/35009  
; CURRENT FILING DATE: 2002-11-01  
; PRIOR APPLICATION NUMBER: US 60/330,855  
; PRIOR FILING DATE: 2001-11-01  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic  
PCT-US02-35009-11

Query Match 50.0%; Score 18; DB 1; Length 7;  
Best Local Similarity 56.7%; Pred. No. 4e+05;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 NVLQ 8  
| | | |  
Db 1 NVLQ 6

RESULT 14

US-09-637-216B-33  
; Sequence 33, Application US/09637216B  
; GENERAL INFORMATION:  
; APPLICANT: WASHINGTON UNIVERSITY  
; TITLE OF INVENTION: ANTI-BACTERIAL COMPOUNDS DIRECTED AGAINST PILUS  
; TITLE OF INVENTION: BIOGENESIS, ADHESION AND ACTIVITY; CO-CRYSTALS OF PILUS  
; TITLE OF INVENTION: SUBUNITS AND METHODS OF USE THEREOF  
; FILE REFERENCE: WSHU2005.1  
; CURRENT APPLICATION NUMBER: US/09/637,216B  
; CURRENT FILING DATE: 2000-08-11  
; PRIOR APPLICATION NUMBER: US 60/148,280  
; PRIOR FILING DATE: 1999-08-11  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 33  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthesized  
; OTHER INFORMATION: Sequence  
US-09-637-216B-33

Query Match 50.0%; Score 18; DB 5; Length 7;  
Best Local Similarity 75.0%; Pred. No. 4e+05;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLQ 6  
| | | |  
Db 1 NVLQ 4



RESULT 15  
US-10-284-252-11  
; Sequence 11, Application US/10284252  
; GENERAL INFORMATION:  
; APPLICANT: Jones, C. Hal  
; APPLICANT: Dexter, Paul L.  
; APPLICANT: Evans, Amy K.  
; APPLICANT: Hruby, Dennis E.  
; TITLE OF INVENTION: DRGP Protease: Cleavage Site  
; TITLE OF INVENTION: Identification and Proteolysis of a Natural Target in E.  
; TITLE OF INVENTION: Coll  
; FILE REFERENCE: 016921-169  
; CURRENT APPLICATION NUMBER: US/10/284,252  
; CURRENT FILING DATE: 2002-10-31  
; PRIOR APPLICATION NUMBER: US 60/330,855  
; PRIOR FILING DATE: 2001-11-01  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic  
US-10-284-252-11

Query Match 50.0%; Score 18; DB 6; Length 7;  
Best Local Similarity 66.7%; Pred. No. 4e+05;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 NVLQXT 8  
   | | | |  
Db 1 NVLHYT 6

Search completed: February 26, 2003, 15:52:27  
Job time : 23 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:42:51 ; Search time 14 Seconds  
(without alignments)  
16.813 Million cell updates/sec

Title: US-09-476-485A-24

Perfect score: 36

Sequence: 1 TNNVLOXT 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 53074

Minimum DB seq length: 0

Maximum DB seq length: 8

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA.\*
- 1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*
  - 3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep.\*
  - 4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep.\*
  - 6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 34    | 94.4        | 8      | 4     | US-08-881-189B-24 |
| 2          | 20    | 55.6        | 7      | 2     | US-08-702-105A-22 |
| 3          | 20    | 55.6        | 7      | 3     | US-08-702-110A-22 |
| 4          | 20    | 55.6        | 7      | 4     | US-09-325-571-22  |
| 5          | 18    | 50.0        | 6      | 2     | US-07-662-764D-25 |
| 6          | 17    | 47.2        | 5      | 3     | US-08-915-189-36  |
| 7          | 17    | 47.2        | 5      | 3     | US-08-915-189-37  |
| 8          | 17    | 47.2        | 5      | 3     | US-08-915-189-38  |
| 9          | 17    | 47.2        | 5      | 3     | US-08-972-760-36  |
| 10         | 17    | 47.2        | 5      | 3     | US-08-972-760-37  |
| 11         | 17    | 47.2        | 5      | 3     | US-08-972-760-38  |
| 12         | 17    | 47.2        | 5      | 4     | US-09-089-645A-36 |
| 13         | 17    | 47.2        | 5      | 4     | US-09-089-645A-37 |
| 14         | 17    | 47.2        | 5      | 4     | US-09-089-645A-38 |
| 15         | 17    | 47.2        | 5      | 4     | US-09-503-998-36  |
| 16         | 17    | 47.2        | 5      | 4     | US-09-503-998-37  |
| 17         | 17    | 47.2        | 5      | 4     | US-09-503-998-38  |
| 18         | 17    | 47.2        | 7      | 1     | US-08-346-333-32  |
| 19         | 17    | 47.2        | 7      | 2     | US-08-352-973A-34 |
| 20         | 17    | 47.2        | 7      | 2     | US-08-946-241B-11 |
| 21         | 17    | 47.2        | 7      | 3     | US-09-309-053-11  |
| 22         | 17    | 47.2        | 7      | 4     | US-09-089-821-18  |
| 23         | 17    | 47.2        | 7      | 5     | PCT-US91-07506-32 |
| 24         | 17    | 47.2        | 7      | 6     | 5210075-58        |
| 25         | 17    | 47.2        | 8      | 1     | US-08-165-038-25  |
| 26         | 17    | 47.2        | 8      | 2     | US-08-876-781-25  |
| 27         | 17    | 47.2        | 8      | 2     | US-09-016-366A-65 |

|    |    |      |   |   |                   |                   |
|----|----|------|---|---|-------------------|-------------------|
| 28 | 16 | 44.4 | 4 | 1 | US-08-456-424-48  | Sequence 48, Appl |
| 29 | 16 | 44.4 | 5 | 4 | US-08-974-549A-81 | Sequence 81, Appl |
| 30 | 16 | 44.4 | 5 | 4 | US-08-974-549A-97 | Sequence 97, Appl |
| 31 | 16 | 44.4 | 6 | 1 | US-08-486-721A-19 | Sequence 19, Appl |
| 32 | 16 | 44.4 | 7 | 1 | US-07-968-781A-52 | Sequence 52, Appl |
| 33 | 16 | 44.4 | 7 | 1 | US-08-486-721A-5  | Sequence 5, Appl  |
| 34 | 16 | 44.4 | 7 | 1 | US-08-486-721A-10 | Sequence 10, Appl |
| 35 | 16 | 44.4 | 7 | 1 | US-08-486-721A-12 | Sequence 12, Appl |
| 36 | 16 | 44.4 | 7 | 3 | US-08-874-678-11  | Sequence 11, Appl |
| 37 | 16 | 44.4 | 7 | 3 | US-08-643-839-11  | Sequence 11, Appl |
| 38 | 16 | 44.4 | 7 | 4 | US-09-410-025-11  | Sequence 11, Appl |
| 39 | 16 | 44.4 | 7 | 4 | US-09-348-886-11  | Sequence 11, Appl |
| 40 | 16 | 44.4 | 8 | 1 | US-08-279-906A-12 | Sequence 12, Appl |
| 41 | 16 | 44.4 | 8 | 1 | US-08-279-906A-13 | Sequence 13, Appl |
| 42 | 16 | 44.4 | 8 | 1 | US-08-486-721A-6  | Sequence 6, Appl  |
| 43 | 16 | 44.4 | 8 | 4 | US-08-953-033-22  | Sequence 22, Appl |
| 44 | 16 | 44.4 | 8 | 4 | US-09-139-802-157 | Sequence 157, App |
| 45 | 15 | 41.7 | 4 | 4 | US-09-177-249-231 | Sequence 231, App |

## ALIGNMENTS

RESULT 1  
US-08-881-189B-24  
; Sequence 24, Application US/08881189B  
; Patent No. 6310195  
; GENERAL INFORMATION:  
; APPLICANT: Colucci et al.  
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED  
; PROGENITOR CELL PRESERVATION FACTOR  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron, LLP  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11753

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

COMPUTER: IBM compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: WordPerfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/881,189B

FILING DATE: June 24, 1997

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Felt, Irving N.

REGISTRATION NUMBER: 28,601

REFERENCE/DOCKET NUMBER: 381-44 PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 822-3550

TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 8 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-881-189B-24

Query Match 94.4%; Score 34; DB 4; Length 8;  
Best Local Similarity 87.5%; Pred. No. 2e+05;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLOXT 8

Db 1 TNNVLOXT 8

```

; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner L.L.P.
; STREET: 1300 I Street N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,110A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/874,503
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32984
; REFERENCE/DOCKET NUMBER: 05387.0056-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-702-110A-22

Query Match 55.6%; Score 20; DB 3; Length 7;
Best Local Similarity 75.0%; Pred. NO. 2e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4
Db 4 TNNI 7

RESULT 4
US-09-325-571-22
; Sequence 22, Application US/09325571
; Patent No. 6261559
; GENERAL INFORMATION:
; APPLICANT: Levitt, Roy C.
; APPLICANT: Maloy, W. Lee
; APPLICANT: Kari, U. Prasad
; APPLICANT: Nicolaides, Nicholas C.
; TITLE OF INVENTION: Asthma Associated Factors As Targets For
; TITLE OF INVENTION: Treating Atopic Allergies Including Asthma And Related
; TITLE OF INVENTION: Disorders
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner L.L.P.
; STREET: 1300 I Street N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/325,571
; FILING DATE:

```

```
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/874,503
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32984
; REFERENCE/DOCKET NUMBER: 05387.0056-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-325-571-22

Query Match 55.6%; Score 20; DB 4; Length 7;
Best Local Similarity 75.0%; Pred. No. 2e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0;

Qy 1 TNNV 4
Db 4 TNNI 7

RESULT 5
US-07-662-764D-25
; Sequence 25, Application US/07662764D
; Patent No. 5866363
; GENERAL INFORMATION:
; APPLICANT: Piecznik, George
; TITLE OF INVENTION: METHOD AND MEANS FOR SORTING AND
; TITLE OF INVENTION: IDENTIFYING BIOLOGICAL INFORMATION
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLZ &
; ADDRESSEE: MENTLIK
; STREET: 600 South, Avenue West
; CITY: Westfield
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07090
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/662,764D
; FILING DATE: 28-FEB-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/201,358
; FILING DATE: 26-MAY-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/770,390
; FILING DATE: 28-AUG-1985
; ATTORNEY/AGENT INFORMATION:
; NAME: Foley, Shawn P.
; REGISTRATION NUMBER: 33,071
; REFERENCE/DOCKET NUMBER: ICTECH/0002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-654-5000
; TELEFAX: 908-654-7866
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: protein
; US-07-662-764D-25

Query Match 50.0%; Score 18; DB 2; Length 6;
Best Local Similarity 75.0%; Pred. No. 2e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0;

Qy 3 NVLQ 6
Db 1 NILO 4

RESULT 6
US-08-915-189-36
; Sequence 36, Application US/08915189
; Patent No. 6001965
; Patent No. 6001965 6001955
; GENERAL INFORMATION:
; APPLICANT: Livant, Donna L
; TITLE OF INVENTION: Anticancer Compounds and Methods
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medien & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,189
; FILING DATE: 20-AUG-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: UM-02877
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; US-08-915-189-36

Query Match 47.2%; Score 17; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;

Qy 1 TNN 3
Db 3 TNN 5

RESULT 7
US-08-915-189-37
; Sequence 37, Application US/08915189
; Patent No. 6001965
; Patent No. 6001965 6001955
; GENERAL INFORMATION:
; APPLICANT: Livant, Donna L
; TITLE OF INVENTION: Anticancer Compounds and Methods
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medien & Carroll, LLP
```

; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States Of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/915,189  
; FILING DATE: 20-AUG-1997  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Peter G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: UM-02877  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 37:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
US-08-915-189-37

Query Match 47.2%; Score 17; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNN 3  
Db 3 TNN 5

RESULT 8  
US-08-915-189-38  
; Sequence 38, Application US/08915189  
; Patent No. 6001965  
; Patent No. 6001965 6001955  
; GENERAL INFORMATION:  
; APPLICANT: Livant, Donna L  
; TITLE OF INVENTION: Anticancer Compounds and Methods  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States Of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/915,189  
; FILING DATE: 20-AUG-1997  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Peter G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: UM-02877  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 38:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 5 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
US-08-915-189-38

Query Match 47.2%; Score 17; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNN 3  
Db 3 TNN 5

RESULT 9  
US-08-972-760-36  
; Sequence 36, Application US/08972760  
; Patent No. 6025150  
; GENERAL INFORMATION:  
; APPLICANT: Livant, Donna L  
; TITLE OF INVENTION: Methods and Compositions for Wound  
; TITLE OF INVENTION: Healing  
; NUMBER OF SEQUENCES: 85  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States Of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/972,760  
; FILING DATE: 18-NOV-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/754,322  
; FILING DATE: 21-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Peter G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: UM-03057  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
US-08-972-760-36

Query Match 47.2%; Score 17; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNN 3  
Db 3 TNN 5

RESULT 10  
US-08-972-760-37  
; Sequence 37, Application US/08972760  
; Patent No. 6025150

;; GENERAL INFORMATION:  
;; APPLICANT: Livant, Donna L  
;; TITLE OF INVENTION: Methods and Compositions for Wound  
;; TITLE OF INVENTION: Healing  
;; NUMBER OF SEQUENCES: 85  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Medlen & Carroll, LLP  
;; STREET: 220 Montgomery Street, Suite 2200  
;; CITY: San Francisco  
;; STATE: California  
;; COUNTRY: United States Of America  
;; ZIP: 94104  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/972,760  
;; FILING DATE: 18-NOV-1997  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/754,322  
;; FILING DATE: 21-NOV-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Carroll, Peter G.  
;; REGISTRATION NUMBER: 32,837  
;; REFERENCE/DOCKET NUMBER: UM-03057  
;; TELEPHONE: (415) 705-8410  
;; TELEFAX: (415) 397-8338  
;; INFORMATION FOR SEQ ID NO: 37:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 5 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: not relevant  
;; TOPOLOGY: not relevant  
;; MOLECULE TYPE: peptide  
;; US-08-972-760-37

Query Match 47.2%; Score 17; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0;

QY 1 TNN 3  
Db 3 TNN 5

RESULT 11  
US-08-972-760-38  
; Sequence 38, Application US/08972760  
; Patent No. 6025150  
; GENERAL INFORMATION:  
; APPLICANT: Livant, Donna L  
; TITLE OF INVENTION: Methods and Compositions for Wound  
; TITLE OF INVENTION: Healing  
; NUMBER OF SEQUENCES: 85  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States Of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/972,760  
; FILING DATE: 18-NOV-1997

;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/754,322  
;; FILING DATE: 21-NOV-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Carroll, Peter G.  
;; REGISTRATION NUMBER: 32,837  
;; REFERENCE/DOCKET NUMBER: UM-03057  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 705-8410  
;; TELEFAX: (415) 397-8338  
;; INFORMATION FOR SEQ ID NO: 38:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 5 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: not relevant  
;; TOPOLOGY: not relevant  
;; MOLECULE TYPE: peptide  
;; US-08-972-760-38

Query Match 47.2%; Score 17; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0;

QY 1 TNN 3  
Db 3 TNN 5

RESULT 12  
US-09-089-645A-36  
; Sequence 36, Application US/09089645A  
; Patent No. 6140068  
; GENERAL INFORMATION:  
; APPLICANT: Livant, Donna L  
; TITLE OF INVENTION: Protease Resistant Compositions for  
; TITLE OF INVENTION: Wound Healing  
; NUMBER OF SEQUENCES: 86  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States Of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/089,645A  
; FILING DATE: 03-JUN-1998  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/754,322  
; FILING DATE: 21-NOV-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/972,760  
; FILING DATE: 18-NOV-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Peter G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: UM-03349  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant

; MOLECULE TYPE: peptide  
US-09-089-645A-36

Query Match 47.2%; Score 17; DB 4; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0;

Qy 1 TNN 3  
|||  
Db 3 TNN 5

## RESULT 13

US-09-089-645A-37  
; Sequence 37, Application US/09089645A  
; Patent No. 6140068  
; GENERAL INFORMATION:  
; APPLICANT: Livant, Donna L  
; TITLE OF INVENTION: Protease Resistant Compositions for  
; Wound Healing  
; NUMBER OF SEQUENCES: 86  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States Of America  
; ZIP: 94104

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/089,645A  
; FILING DATE: 03-JUN-1998  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/754,322  
; FILING DATE: 21-NOV-1996

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/972,760  
; FILING DATE: 18-NOV-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Peter G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: UM-03349  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 37:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
US-09-089-645A-37

Query Match 47.2%; Score 17; DB 4; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0;

Qy 1 TNN 3  
|||  
Db 3 TNN 5

## RESULT 14

US-09-089-645A-38  
; Sequence 38, Application US/09089645A  
; Patent No. 6140068  
; GENERAL INFORMATION:

; APPLICANT: Livant, Donna L  
; TITLE OF INVENTION: Protease Resistant Compositions for  
; Wound Healing  
; NUMBER OF SEQUENCES: 86  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States Of America  
; ZIP: 94104

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/089,645A  
; FILING DATE: 03-JUN-1998  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/754,322  
; FILING DATE: 21-NOV-1996

; PRIOR APPLICATION DATA: US 08/972,760  
; APPLICATION NUMBER: 18-NOV-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Peter G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: UM-03349  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 38:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
US-09-089-645A-38

Query Match 47.2%; Score 17; DB 4; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0;

Qy 1 TNN 3  
|||  
Db 3 TNN 5

## RESULT 15

US-09-503-998-36  
; Sequence 36, Application US/09503998  
; Patent No. 6331409  
; GENERAL INFORMATION:  
; APPLICANT: Livant, Donna L  
; TITLE OF INVENTION: Methods and Compositions for Wound  
; Healing  
; NUMBER OF SEQUENCES: 85  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States Of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:



APPLICATION NUMBER: US/09/503,998  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/972,760  
FILING DATE: 18-NOV-1997  
APPLICATION NUMBER: US 08/754,322  
FILING DATE: 21-NOV-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Peter G.  
REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: UM-03057  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
STRANDEDNESS: No. 6331409 Relevant  
TOPOLOGY: No. 6331409 Relevant  
MOLECULE TYPE: peptide  
US-09-503-998-36

Query Match 47.2%; Score 17; DB 4; Length 5;  
Best Local Similarity 100.0%; Pred. No. 2e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNN 3  
Db 3 TNN 5

Search completed: February 26, 2003, 15:46:03  
Job time : 15 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:55:30 ; Search time 45 Seconds  
(without alignments)  
29.908 Million cell updates/sec

Title: US-09-476-485A-31

Perfect score: 66

Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 2093

Minimum DB seq length: 0

Maximum DB seq length: 14

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 27    | 40.9        | 14     | 2     | PA0007      |
| 2          | 25    | 37.9        | 12     | 2     | B61497      |
| 3          | 22    | 33.3        | 10     | 2     | C39572      |
| 4          | 22    | 33.3        | 14     | 2     | C59137      |
| 5          | 21    | 31.8        | 14     | 2     | F61497      |
| 6          | 21    | 31.8        | 14     | 2     | A35105      |
| 7          | 21    | 31.8        | 14     | 2     | B61597      |
| 8          | 20    | 30.3        | 12     | 2     | S26547      |
| 9          | 19    | 28.8        | 14     | 2     | B44854      |
| 10         | 18    | 27.3        | 9      | 2     | PD0443      |
| 11         | 18    | 27.3        | 10     | 2     | S38305      |
| 12         | 18    | 27.3        | 12     | 2     | A61360      |
| 13         | 18    | 27.3        | 13     | 2     | PQ0445      |
| 14         | 18    | 27.3        | 13     | 2     | PS0443      |
| 15         | 18    | 27.3        | 14     | 2     | PT0077      |
| 16         | 18    | 27.3        | 14     | 2     | S07768      |
| 17         | 17.5  | 26.5        | 13     | 2     | S23640      |
| 18         | 17    | 25.8        | 10     | 1     | ECU04M      |
| 19         | 17    | 25.8        | 10     | 2     | C60788      |
| 20         | 17    | 25.8        | 10     | 2     | A60787      |
| 21         | 17    | 25.8        | 10     | 2     | A60527      |
| 22         | 17    | 25.8        | 10     | 2     | I60527      |
| 23         | 17    | 25.8        | 10     | 2     | A24867      |
| 24         | 17    | 25.8        | 10     | 2     | D61440      |
| 25         | 17    | 25.8        | 13     | 2     | A33660      |
| 26         | 17    | 25.8        | 13     | 2     | S10562      |
| 27         | 17    | 25.8        | 13     | 2     | I77387      |
| 28         | 17    | 25.8        | 14     | 2     | S39932      |
| 29         | 17    | 25.8        | 14     | 2     | B28018      |

## ALIGNMENTS

## RESULT 1

PA0007

lectin B1 - Psophocarpus scandens (fragment)

C:Species: Psophocarpus scandens

C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Mar-1995

C:Accession: PA0007

R:Kortt, A.A.

Phytochemistry 27, 2847-2855, 1988

A:Title: Isolation and characterization of the lectins from the seeds of Psophocarpus

A:Reference number: PA0005

A:Accession: PA0007

A:Molecule type: protein

A:Residues: 1-14 <KOR>

A:Experimental source: seed

C:Comment: The seeds of Psophocarpus contain two distinct groups of lectins which can

C:Keywords: lectin

Query Match 40.9%; Score 27; DB 2; Length 14;

Best Local Similarity 40.0%; Pred. No. 64;

Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Oy 2 QSLSFXTKFK 11

Db 1 ETISFNFNQF 10

## RESULT 2

B61497

seed protein ws-17 - winged bean (fragment)

C:Species: Psophocarpus tetragonolobus (winged bean)

C:Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 07-Oct-1994

C:Accession: B61497

R:Hirano, H.

J. Protein Chem. 8, 115-130, 1989

A:Title: Microsequence analysis of winged bean seed proteins electrophoretically from two-

A:Reference number: A61491; PMID:89351606; PMID:2765119

A:Accession: B61497

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-12 <HIR>

C:Keywords: seed

Query Match 37.9%; Score 25; DB 2; Length 12;

Best Local Similarity 44.4%; Pred. No. 1.3e+02;

Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 3 SLSFXFTKFK 11

Db 2 TISFNFNQF 10

## RESULT 3

```

RESULT 6
A35105
hypoetical protein - Neurospora crassa mitochondrion (fragment)
C:Species: mitochondrion Neurospora crassa
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 07-Dec-1999
C:Accession: A35105
R:Saville, B.J.; Collins, R.A.
Cell 61, 685-696, 1990
A:Title: A site-specific self-cleavage reaction performed by a novel RNA in neurospor
A:Reference number: A35105; MUID:90263093; PMID:2160856
A:Accession: A35105
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-14 <SAV>
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC3
C:Keywords: mitochondrion

Query Match      31.8%; Score 21; DB 2; Length 14;
Best Local Similarity 50.0%; Pred. No. 8.8e+02;
Matches      5; Conservative      1; Mismatches      4; Indels      0; Gaps      0;

QY      4 LSFYFTKFDL 13
      ||| : | |
Db      1 LSLFWTLQL 10

RESULT 7
B61597
cytochrome P450 AL-2 - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 06-Jan-1995
C:Accession: B61597
R:Shimeno, H.; Toda, S.; Nagamatsu, A.
Drug Metab. Dispos. 19, 291-297, 1991
A:Title: Purification and aminopyrine monooxygenase activity of liver microsomal cyto
A:Reference number: A61597; MUID:91292910; PMID:1676625

```

```

A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-14 <SH1>

Query Match      31.8%; Score 21; DB 2; Length 14;
Best Local Similarity 55.6%; Pred. No. 8.8e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      3 SLSFXFTKF 11
        |||||
Db       6 SLSFLLVGF 14

RESULT 8
S26547
T-cell receptor beta chain (clone Cw3/10.1) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 17-Apr-1998 #text_change 17-Mar-1999
C:Accession: S26547
R:Casanova, J.L.; Cerottini, J.C.; Matthes, M.; Necker, A.; Gournier, H.; Barra, C.;
J. Exp. Med. 176; 439-447, 1992
A:Title: H-2-restricted cytolytic T lymphocytes specific for HLA display T cell receptor
A:Reference number: S26512; MUID:92364346; PMID:1380061
A:Accession: S26547
A:Molecule type: mRNA
A:Residues: 1-12 <CAS>
A:Cross-references: EMBL:X67997
A:Experimental source: cytolytic T-lymphocyte, clone Cw3/10.1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor

Query Match      30.3%; Score 20; DB 2; Length 12;

```

Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 AQSLSFXFT 9  
 | | | | |  
 Db 2 ASSTGFDYT 10

## RESULT 9

B44854  
 L-2,4-diaminobutyrate decarboxylase (EC 4.1.1.-) - Vibrio alginolyticus (fragment)  
 C:Species: Vibrio alginolyticus  
 C:Date: 31-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 10-Nov-1995  
 C:Accession: B44854; B41817  
 R:Yamamoto, S.; Tsuzaki, Y.; Tougou, K.; Shinoda, S.  
 J. Gen. Microbiol. 138, 1461-1465, 1992  
 A:Title: Purification and characterization of L-2,4-diaminobutyrate decarboxylase from  
 A:Reference number: A44854; MUID:92381494; PMID:1512577  
 A:Accession: B44854  
 A:Molecule type: protein  
 A:Residues: 1-14 <YAM>  
 A:Note: sequence extracted from NCBI backbone (NCBIP:112332)  
 C:Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 28.8%; Score 19; DB 2; Length 14;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+03;  
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 9 TKFDLD 14  
 | | | | |  
 Db 2 TAFEVD 7

## RESULT 10

PD0443  
 3-oxoacid CoA-transferase (EC 2.8.3.5) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 05-Feb-1999  
 C:Accession: PD0443  
 R:Kawakami, T.; Uchida, T.; Sakai, T.; Kamo, M.; Morimasa, T.; Tsugita, A.  
 Submitted to JIPID, August 1998  
 A:Description: Proteome analysis of mouse brain.  
 A:Reference number: PD0441  
 A:Contents: Striatum  
 A:Accession: PD0443  
 A:Molecule type: protein  
 A:Residues: 1-9 <KAM>  
 C:Keywords: CoA-transferase

Query Match 27.3%; Score 18; DB 2; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 2.8e+05;  
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 TKFDLD 14  
 | | | | |  
 Db 1 TKFYTD 6

## RESULT 11

S38305  
 lectin GNL2 alpha chain - kidney bean (fragment)  
 C:Species: Phaseolus vulgaris (kidney bean)  
 C:Date: 19-May-1994 #sequence\_revision 27-Feb-1997 #text\_change 13-Mar-1997  
 C:Accession: S38305  
 R:Kamenura, K.; Furuichi, Y.; Umekawa, H.; Takahashi, T.  
 Biochim. Biophys. Acta 1158, 181-188, 1993  
 A:Title: Purification and characterization of novel lectins from Great Northern bean, Ph  
 A:Reference number: S38304; MUID:94002183; PMID:8399319  
 A:Accession: S38305  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-10 <KAM>

Query Match 27.3%; Score 18; DB 2; Length 10;

Best Local Similarity 50.0%; Pred. No. 2.3e+03;  
 Matches 4; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AQSLSFXF 8  
 | | | | |  
 Db 1 ATETSFSS 8

## RESULT 12

A61360  
 vespakinin M - hornet (Vespa mandarinia)  
 C:Species: Vespa mandarinia  
 C:Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 18-Aug-2000  
 C:Accession: A61360  
 R:Kishimura, H.; Yasuhara, T.; Yoshida, H.; Nakajima, T.  
 Chem. Pharm. Bull. 24, 2896-2897, 1976  
 A:Title: Vespakinin-M, a novel bradykinin analogue containing hydroxyproline, in the  
 A:Reference number: A61360; MUID:77114342; PMID:1017116  
 A:Accession: A61360  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-12 <KIS>  
 C:Superfamily: unassigned animal peptides  
 C:Keywords: hydroxyproline; venom  
 F:4/Modified site: 4-hydroxyproline (pro) #status experimental

Query Match 27.3%; Score 18; DB 2; Length 12;  
 Best Local Similarity 42.9%; Pred. No. 2.8e+03;  
 Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 FTRFDLD 14  
 | | | | |  
 Db 6 FSPFRID 12

## RESULT 13

PQ0445  
 urotensin II - laughing frog  
 C:Species: Rana ridibunda (laughing frog)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 03-Mar-1995  
 C:Accession: PQ0445  
 R:Conlon, J.M.; O'Harte, F.; Smith, D.D.; Tonon, M.C.; Vaudry, H.  
 Biochem. Biophys. Res. Commun. 188, 578-583, 1992  
 A:Title: Isolation and primary structure of urotensin II from the brain of a tetrapod  
 A:Reference number: PQ0445; MUID:93075134; PMID:1445302  
 A:Accession: PQ0445  
 A:Molecule type: protein  
 A:Residues: 1-13 <CON>  
 A:Experimental source: brain  
 C:Superfamily: urotensin II

Query Match 27.3%; Score 18; DB 2; Length 13;  
 Best Local Similarity 45.5%; Pred. No. 3e+03;  
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKF 11  
 | | | | |  
 Db 1 AGNLSECFWKY 11

## RESULT 14

PS0443  
 potassium channel protein slo G3 - fruit fly (Drosophila melanogaster) (fragment)  
 C:Species: Drosophila melanogaster  
 C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 16-Feb-1997  
 C:Accession: PS0443  
 R:Adelman, J.P.; Shen, K.Z.; Kavanaugh, M.P.; Warren, R.A.; Wu, Y.N.; Lagrutta, A.; E  
 Neuron 9, 209-216, 1992  
 A:Title: Calcium-activated potassium channels expressed from cloned complementary DNA  
 A:Reference number: JH0697; MUID:92360298; PMID:1497890  
 A:Accession: PS0443  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA

A;Residues: 1-13 <ADE>  
C;Comment: This potassium channel is activated by calcium.  
C;Genetics:  
A;Gene: FlyBase:slo  
A;Cross-references: FlyBase:FBgn0003429  
C;Keywords: alternative splicing; ion channel; potassium channel; transmembrane protein

Query Match 27.3%; Score 18; DB 2; Length 13;  
Best Local Similarity 75.0%; Pred. No. 3e+03;  
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 9 TKFD 12  
: : :  
Db 7 SKFD 10

RESULT 15  
PT0077  
proteochondroitin core protein - rat (fragment)  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 04-Sep-1998 #sequence\_revision 04-Sep-1998 #text\_change 04-Sep-1998  
C;Accession: PT0077  
R;Marcum, J.A.; Thompson, M.A.  
Biochem. Biophys. Res. Commun. 175, 706-712, 1991  
A;Title: The amino-terminal region of a proteochondroitin core protein, secreted by aorta from human bone.  
A;Reference number: PT0077; PMID:91207372; PMID:2018513  
A;Accession: PT0077  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-14 <MAR>  
C;Superfamily: decorin; leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan

Query Match 27.3%; Score 18; DB 2; Length 14;  
Best Local Similarity 57.1%; Pred. No. 3.2e+03;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 8 FTKFDLD 14  
: : : :  
Db 8 FWDFDLD 14

Search completed: February 26, 2003, 14:58:06  
Job time : 45 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:51:55 ; Search time 10 Seconds  
(without alignments)  
58.067 Million cell updates/sec

Title: US-09-476-485A-31  
Perfect score: 66  
Sequence: 1 AQLSLFXFTKFDLD 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 633

Minimum DB seq length: 0  
Maximum DB seq length: 14

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 27    | 40.9        | 14     | 1     | LECB_PSOSC  |
| 2          | 19    | 28.8        | 13     | 1     | ODPA_CANFA  |
| 3          | 18    | 27.3        | 10     | 1     | UXA6_CHLTR  |
| 4          | 18    | 27.3        | 13     | 1     | TEML_RANTE  |
| 5          | 18    | 27.3        | 14     | 1     | DHSL_ANACY  |
| 6          | 17    | 25.8        | 10     | 1     | LCMS_LEUMA  |
| 7          | 17    | 25.8        | 10     | 1     | TKL4_LOCM1  |
| 8          | 17    | 25.8        | 10     | 1     | TKN1_SCYCA  |
| 9          | 16    | 24.2        | 8      | 1     | ACT_CARNA   |
| 10         | 16    | 24.2        | 10     | 1     | FARP_MANSE  |
| 11         | 16    | 24.2        | 11     | 1     | BRK_MEGFL   |
| 12         | 16    | 24.2        | 11     | 1     | MLG_THETS   |
| 13         | 15    | 22.7        | 10     | 1     | TRP9_LEUMA  |
| 14         | 15    | 22.7        | 12     | 1     | V25K_WSSV   |
| 15         | 14    | 21.2        | 9      | 1     | FAR3_CALVO  |
| 16         | 14    | 21.2        | 9      | 1     | UHA2_HUMAN  |
| 17         | 14    | 21.2        | 10     | 1     | FAR6_PANRE  |
| 18         | 14    | 21.2        | 10     | 1     | MOSQ_CLYJA  |
| 19         | 14    | 21.2        | 12     | 1     | UR2_POLSP   |
| 20         | 14    | 21.2        | 13     | 1     | AU11_LITRA  |
| 21         | 14    | 21.2        | 13     | 1     | AU12_LITRA  |
| 22         | 14    | 21.2        | 13     | 1     | FIBA_CAVPO  |
| 23         | 14    | 21.2        | 13     | 1     | NP2_LYMST   |
| 24         | 14    | 21.2        | 13     | 1     | NP5_LYMST   |
| 25         | 13    | 19.7        | 4      | 1     | PLRE_HIRME  |
| 26         | 13    | 19.7        | 4      | 1     | FMRE_MACNI  |
| 27         | 13    | 19.7        | 7      | 1     | FAR1_MACNI  |
| 28         | 13    | 19.7        | 7      | 1     | FAR1_HELTI  |
| 29         | 13    | 19.7        | 7      | 1     | FAR1_PROCL  |
| 30         | 13    | 19.7        | 7      | 1     | FAR2_PROCL  |
| 31         | 13    | 19.7        | 7      | 1     | FAR4_PANRE  |
| 32         | 13    | 19.7        | 7      | 1     | FARB_CALVO  |
| 33         | 13    | 19.7        | 8      | 1     | ALL4_CYDPO  |

|    |    |      |   |   |            |                    |
|----|----|------|---|---|------------|--------------------|
| 34 | 13 | 19.7 | 8 | 1 | FAR1_PANRE | P41872 panagrellus |
| 35 | 13 | 19.7 | 8 | 1 | FAR3_HOMAM | P41486 homarus ame |
| 36 | 13 | 19.7 | 8 | 1 | FAR4_HOMAM | P41487 homarus suu |
| 37 | 13 | 19.7 | 8 | 1 | FAR7_ASCSU | P43171 ascaris suu |
| 38 | 13 | 19.7 | 8 | 1 | FAR8_CALVO | P41863 calliphora  |
| 39 | 13 | 19.7 | 9 | 1 | FAR1_CALVO | P41856 calliphora  |
| 40 | 13 | 19.7 | 9 | 1 | FAR2_CALVO | P41857 calliphora  |
| 41 | 13 | 19.7 | 9 | 1 | FAR2_PANRE | P41873 panagrellus |
| 42 | 13 | 19.7 | 9 | 1 | FAR2_CALVO | P41859 calliphora  |
| 43 | 13 | 19.7 | 9 | 1 | FAR5_ASCSU | P43170 ascaris suu |
| 44 | 13 | 19.7 | 9 | 1 | FAR5_CALVO | P41860 calliphora  |
| 45 | 13 | 19.7 | 9 | 1 | FAR6_CALVO | P41861 calliphora  |

## ALIGNMENTS

RESULT 1  
LECB\_PSOSC  
ID LECB\_PSOSC STANDARD; PRT; 14 AA.  
AC P22584;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE Basic lectin B1 (Fragment).  
OS Psophocarpus scandens.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
OC Psophocarpus.  
OX NCBI\_TaxID=3890;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Seed;  
RA Kortt A.A.;  
RT Isolation and characterization of the lectins from the seeds of  
RT Psophocarpus scandens.;  
RL Phytochemistry 27:2847-2855(1988).  
CC -!- SUBUNIT: DIMER OF TWO IDENTICAL OR NEARLY IDENTICAL SUBUNITS OF  
CC ABOUT 32000 APPARENT MW.  
CC -!- PTM: CONTAINS ABOUT 9% CARBOHYDRATE. HETEROGENEITY IN CARBOHYDRATE  
CC COMPOSITION MAY RESULT IN DIFFERENT ISOLECTIN FORMS.  
CC -!- SIMILARITY: WITH P.TETRAGONOLOBUS BASIC LECTINS IN N-TERMINAL  
CC SEQUENCE, AA COMPOSITION, ERYTHROCYTE AND SUGAR SPECIFICITIES.  
DR PIR; PA0007; PA0007.  
KW Lectin; Glycoprotein.  
FT NON\_TER 14  
SQ SEQUENCE 14 AA; 1732 MW; D804CB43B487C549 CRC64;

Query Match 40.9%; Score 27; DB 1; Length 14;  
Best Local Similarity 40.0%; Pred. No. 26;  
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
QY 2 QLSLFXFTKFK 11  
Db 1 ETISFNQF 10  
:::| | | |

## RESULT 2

ODPA\_CANFA STANDARD; PRT; 13 AA.  
ID ODPA\_CANFA  
AC P49823;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Pyruvate dehydrogenase E1 component alpha subunit, somatic form  
DE (EC 1.2.4.1) (PDH-E1-A type I) (fragment).  
GN PDH-E1.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]

RP SEQUENCE.  
 RC TISSUE=Heart;  
 RX MEDLINE=98163340; PubMed=9504812;  
 RA Dunn M.J., Corbett J.M., Wheeler C.H.;  
 RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of  
 dog heart proteins";  
 RL Electrophoresis 18:2795-2802(1997).  
 CC -!- FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall  
 conversion of pyruvate to acetyl-CoA and CO(2). It contains  
 multiple copies of three enzymatic components: pyruvate  
 dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2) and  
 lipoamide dehydrogenase (E3).  
 CC -!- CATALYTIC ACTIVITY: Pyruvate + lipoamide = S-  
 acetyldihydrolipoamide + CO(2).  
 CC -!- COFACTOR: THIAMINE PYROPHOSPHATE.  
 CC -!- ENZYME REGULATION: E1 ACTIVITY IS REGULATED BY PHOSPHORYLATION  
 (INACTIVATION) AND DEPHOSPHORYLATION (ACTIVATION) OF THE ALPHA  
 SUBUNIT (BY SIMILARITY).  
 CC -!- SUBUNIT: TETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS.  
 CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.  
 DR HSC-2DPAGE: P49823; DOG.  
 KW Glycolysis; Oxidoreductase; Flavoprotein; Thiamine pyrophosphate;  
 KW Mitochondrion; Phosphorylation.  
 FT NON\_TER 13 13  
 SQ SEQUENCE 13 AA; 1510 MW; C97EEBF844095B19 CRC64;  
 Query Match 28.8%; Score 19; DB 1; Length 13;  
 Best Local Similarity 44.4%; Pred. No. 7.8e+02;  
 Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 5 SPXFTKFDL 13  
 Db : : :  
 Db 5 TPEIKKXDL 13

RESULT 3  
 OXAE6\_CHLTR  
 ID UXA6\_CHLTR STANDARD; PRT; 10 AA.  
 AC P38007;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Unknown protein from 2D-page from elementary body (Fragment).  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=L2/434/Bu;  
 RA Binl L., Santucci A., Magi B., Marzocchi B., Sanchez-Campillo M.,  
 RA Comanducci M., Christensen G., Birkelund S., Vretou E., Ratti G.,  
 RA Pallini V.;  
 RL Submitted (SEP-1994) to the SWISS-PROT data bank.  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 PROTEIN IS: 6.44, ITS MW IS: 38.6 kDa.  
 DR Siena-2DPAGE; P38007; -.  
 FT NON\_TER 10 10  
 SQ SEQUENCE 10 AA; 1243 MW; DAD39A33304B5339 CRC64;  
 Query Match 27.3%; Score 18; DB 1; Length 10;  
 Best Local Similarity 33.3%; Pred. No. 9.2e+02;  
 Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 4 LSFXTKFD 12  
 Db : : :  
 Db 1 MNEFKYIKD 9

RESULT 4  
 TEML\_RANTE  
 ID TEML\_RANTE STANDARD; PRT; 13 AA.  
 AC P57104;  
 DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Temporin L.  
 OS Rana temporaria (European common frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.  
 OX NCBI\_TaxID=8407;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Skin;  
 RX MEDLINE=97115050; PubMed=9022710;  
 RA Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,  
 RA Barra D.;  
 RT "Temporins, antimicrobial peptides from the European red frog Rana  
 temporaria";  
 RL Eur. J. Biochem. 242:788-792(1996).  
 CC -!- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST GRAM-NEGATIVE AND  
 GRAM-POSITIVE BACTERIA.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN  
 FAMILY.  
 KW Amphibian skin; Antibiotic; Amidation; Multigene family.  
 FT MOD\_RES 13 13  
 SQ SEQUENCE 13 AA; 1641 MW; 9EBCBIFAF7C325 CRC64;  
 Query Match 27.3%; Score 18; DB 1; Length 13;  
 Best Local Similarity 75.0%; Pred. No. 1.2e+03;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 8 FTKF 11  
 Db : : :  
 Db 5 FSKF 8

RESULT 5  
 DHSL\_ANACY  
 ID DHSL\_ANACY STANDARD; PRT; 14 AA.  
 AC P17874;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Soluble hydrogenase 50 kDa subunit (EC 1.12.-.-) (Fragment).  
 OS Anabaena cylindrica.  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.  
 OX NCBI\_TaxID=1165;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=90126821; PubMed=2129525;  
 RA Ewart G.D., Reed K.C., Smith G.D.;  
 RT "Soluble hydrogenase of Anabaena cylindrica. Cloning and sequencing  
 of a potential gene encoding the tritium exchange subunit.";  
 RL Eur. J. Biochem. 187:215-223(1990).  
 CC -!- FUNCTION: SOLUBLE HYDROGENASE CATALYZES BOTH PRODUCTION AND  
 CONSUMPTION OF HYDROGEN FROM SUITABLE ARTIFICIAL ELECTRON DONORS  
 OR ACCEPTORS. THIS SUBUNIT (50 kDa) IS REQUIRED FOR HYDROGEN  
 PRODUCTION WITH REDUCED METHYL-VIOLOGEN.  
 CC -!- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL SUBUNIT.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 DR PIR; S07768; S07768.  
 KW Oxidoreductase.  
 FT NON\_TER 14 14  
 SQ SEQUENCE 14 AA; 1551 MW; 9254DAFEB141CFF2A CRC64;  
 Query Match 27.3%; Score 18; DB 1; Length 14;  
 Best Local Similarity 60.0%; Pred. No. 1.3e+03;  
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 10 KFDLD 14  
 Db : : :  
 Db 3 EFDYD 7

RESULT 6



LCMS\_LEUMA  
ID LCMS\_LEUMA STANDARD; PRT; 10 AA.  
AC P21144; P41197;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 01-FEB-1994 (Rel. 28, Last annotation update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Leucomyosuppressin (LMS) (Lem-MS).  
OS Leucophaea maderae (Madeira cockroach).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
OC Blaberoidea; Blaberidae; Leucophaea.  
OX NCBI\_TaxID=6988;  
RN [1]  
RP SEQUENCE, AND SYNTHESIS.  
RC TISSUE=Head;  
RX Holman G.M., Cook B.J., Nachman R.J.;  
RT "Isolation, primary structure and synthesis of leucomyosuppressin,  
RT an insect neuropeptide that inhibits spontaneous contractions of the  
RT cockroach hindgut.";  
RL Comp. Biochem. Physiol. 85C:329-333(1986).  
CC -1- FUNCTION: INHIBITS THE SPONTANEOUS CONTRACTIONS OF COCKROACH  
CC PROTODEUM (HINDGUT).  
KW Neuropeptide; Amidation.  
FT MOD\_RES 1 1  
FT MOD\_RES 10 10  
FT MOD\_RES 10 10  
SQ SEQUENCE 10 AA; 1275 MW; D3C45229D2C1EAB2 CRC64;  
Amidation.  
Query Match 25.8%; Score 17; DB 1; Length 10;  
Best Local Similarity 30.0%; Pred. No. 1.4e+03;  
Matches 3; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
QY 2 QSLSFXTKFK 11  
I : I : I  
Db 1 QDVDFVFLRF 10

RESULT 7  
TKL4\_LOCMI  
ID TKL4\_LOCMI STANDARD; PRT; 10 AA.  
AC P30250;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Locustatachykinin IV (TK-IV).  
OS Locusta migratoria (Migratory locust).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;  
OC Acridomorpha; Acridoidea; Acrididae; Oedipodinae; Locusta.  
OX NCBI\_TaxID=7004;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Brain;  
RX MEDLINE=91219696; PubMed=2132575;  
RA Schoofs L., Holman G.M., Hayes T.K., Kochansky J.P., Nachman R.J.,  
RA de Loof A.;  
RT "Locustatachykinin III and IV: two additional insect neuropeptides  
RT with homology to peptides of the vertebrate tachykinin family.";  
RL Regul. Pept. 31:199-212(1990).  
CC -1- FUNCTION: MYOACTIVE PEPTIDE. STIMULATES THE CONTRACTION OF THE  
CC OVIDUCT AND FOREGUT.  
CC OVIDUCT AND FOREGUT.  
CC -1- SIMILARITY: SOME SIMILARITY TO TACHYKININS.  
DR PIR; B60073; ECLQ4M.  
KW Tachykinin; Neuropeptide; Amidation.  
FT MOD\_RES 10 10  
FT MOD\_RES 10 10  
SQ SEQUENCE 10 AA; 1040 MW; 9E52CD71E9C87735 CRC64;  
Query Match 25.8%; Score 17; DB 1; Length 10;  
Best Local Similarity 66.7%; Pred. No. 1.4e+03;  
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AQSLSF 6  
I : I : I  
Db 1 APSLSGF 6

LCMS\_LEUMA  
ID LCMS\_LEUMA STANDARD; PRT; 10 AA.  
AC P21144; P41197;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 01-FEB-1994 (Rel. 28, Last annotation update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Leucomyosuppressin (LMS) (Lem-MS).  
OS Leucophaea maderae (Madeira cockroach).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
OC Blaberoidea; Blaberidae; Leucophaea.  
OX NCBI\_TaxID=6988;  
RN [1]  
RP SEQUENCE, AND SYNTHESIS.  
RC TISSUE=Head;  
RX Holman G.M., Cook B.J., Nachman R.J.;  
RT "Isolation, primary structure and synthesis of leucomyosuppressin,  
RT an insect neuropeptide that inhibits spontaneous contractions of the  
RT cockroach hindgut.";  
RL Comp. Biochem. Physiol. 85C:329-333(1986).  
CC -1- FUNCTION: INHIBITS THE SPONTANEOUS CONTRACTIONS OF COCKROACH  
CC PROTODEUM (HINDGUT).  
KW Neuropeptide; Amidation.  
FT MOD\_RES 1 1  
FT MOD\_RES 10 10  
FT MOD\_RES 10 10  
SQ SEQUENCE 10 AA; 1275 MW; D3C45229D2C1EAB2 CRC64;  
Amidation.  
Query Match 25.8%; Score 17; DB 1; Length 10;  
Best Local Similarity 30.0%; Pred. No. 1.4e+03;  
Matches 3; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
QY 2 QSLSFXTKFK 11  
I : I : I  
Db 1 QDVDFVFLRF 10

RESULT 8  
TKNI\_SCYCA  
ID TKNI\_SCYCA STANDARD; PRT; 10 AA.  
AC P08608;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Scyllorhinin I.  
OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Galeomorphii; Galeoidea; Carchariniiformes;  
OC Scyllorhinidae; Scyllorhinus.  
OX NCBI\_TaxID=7830;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=86192829; PubMed=2422058;  
RA Conlon J.M., Deacon C.F., O'Toole L., Thim L.;  
RT "Scyllorhinin I and II: Two novel tachykinins from dogfish gut.";  
RL FEBS Lett. 200:111-116(1986).  
RN [2]  
RP SEQUENCE.  
RC TISSUE=Brain;  
RX MEDLINE=93292508; PubMed=7685693;  
RA Waugh D., Wang Y., Hazon N., Balmert R.J., Conlon J.M.;  
RT "Primary structures and biological activities of substance-P-related  
RT peptides from the brain of the dogfish, Scyllorhinus canicula.";  
RL Eur. J. Biochem. 214:469-474(1993).  
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
CC MUSCLES.  
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
DR PIR; A24867; A24867.  
DR PIR; S33301; S33301.  
DR InterPro; IPR002040; Tachykinin.  
DR PROSITE; PS00267; TACHYKININ; 1.  
KW Tachykinin; Neuropeptide; Amidation.  
FT MOD\_RES 10 10  
FT MOD\_RES 10 10  
SQ SEQUENCE 10 AA; 1219 MW; D0602D6B59C33AA9 CRC64;  
Query Match 25.8%; Score 17; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 10 KFD 12  
I : I : I  
Db 2 KFD 4

RESULT 9  
ACT\_CARMA  
ID ACT\_CARMA STANDARD; PRT; 8 AA.  
AC P80709;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Actin (Fragment).  
OS Carcinus maenas (Common shore crab) (Green crab).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;  
OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;  
OC Brachyura; Eubrachyura; Portunoidae; Portunidae; Carcinus.  
OX NCBI\_TaxID=6759;  
RN [1]  
RP SEQUENCE.  
RA Lachaise F., Somme G., Carpentier G., Granjeon E., Webster S.,  
RA Baghdassarian D.;  
RT "A transaldolase. An enzyme implicated in crab steroidogenesis.";  
RL Endocrine 5:23-32(1996).  
CC -1- FUNCTION: ACTINS ARE HIGHLY CONSERVED PROTEINS THAT ARE INVOLVED  
CC IN VARIOUS TYPES OF CELL MOTILITY AND ARE UBIQUITOUSLY EXPRESSED  
CC IN ALL EUKARYOTIC CELLS.



Db 6 FRWDKF 11

## RESULT 13

TRP9\_LEUMA  
ID TRP9\_LEUMA STANDARD; PRT; 10 AA.  
AC P81741;  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Tachykinin-related peptide 9 (LemRP 9).  
OS Leucophaea maderae (Madeira cockroach).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;  
OC Blaberoidea; Blaberidae; Leucophaea.  
OX NCBI\_TaxID=6988;  
RN [1]

## SEQUENCE, AND MASS SPECTROMETRY.

RC TISSUE=Brain;  
RX MEDLINE=97269266; PubMed=9114447;  
RA Muren J.E., Naessel D.R.;  
RT "Seven tachykinin-related peptides isolated from the brain of the  
RT Madeira cockroach; evidence for tissue-specific expression of  
RT isoforms.";  
RL Peptides 18:7-15(1997).  
CC -!- FUNCTION: MYOACTIVE PEPTIDE. INCREASES THE AMPLITUDE AND FREQUENCY  
CC OF SPONTANEOUS CONTRACTIONS AND TONUS OF HINDGUT MUSCLE.  
CC -!- TISSUE SPECIFICITY: BRAIN.  
CC -!- MASS SPECTROMETRY: MW=1081.5; METHOD=MALDI.  
CC -!- SIMILARITY: SOME SIMILARITY TO TACHYKININS.  
KW Tachykinin; Neuropeptide; Amidation.  
FT MOD\_RES 10 10  
SQ SEQUENCE 10 AA; 1081 MW; 9E469D66D9C87685 CRC64;

Query Match 22.7%; Score 15; DB 1; Length 10;  
Best Local Similarity 50.0%; Pred. No. 3.4e+03;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSF 6  
| | | |  
Db 1 APSMGF 6

## RESULT 14

V25K\_WSSV  
ID V25K\_WSSV STANDARD; PRT; 12 AA.  
AC P82004;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 25 kDa structural polyprotein (Fragment).  
OS White spot syndrome virus (WSSV).  
OC Viruses; dsDNA viruses, no RNA stage; Nimaviridae.  
OX NCBI\_TaxID=92652;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=South Carolina;  
RX MEDLINE=20214217; PubMed=10752552;  
RA Wang Q., Poulos B.T., Lightner D.V.;  
RT "Protein analysis of geographic isolates of shrimp white spot syndrome  
RT virus.";  
RL Arch. Virol. 145:263-274(2000).  
CC -!- FUNCTION: STRUCTURAL COMPONENT OF THE VIRION.  
FT NON\_TER 12 12  
SQ SEQUENCE 12 AA; 1283 MW; C5409AD9ECB731A9 CRC64;

Query Match 22.7%; Score 15; DB 1; Length 12;  
Best Local Similarity 50.0%; Pred. No. 4.1e+03;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 LSFXT 9  
| | | |

Db 3 LSFTLS 8

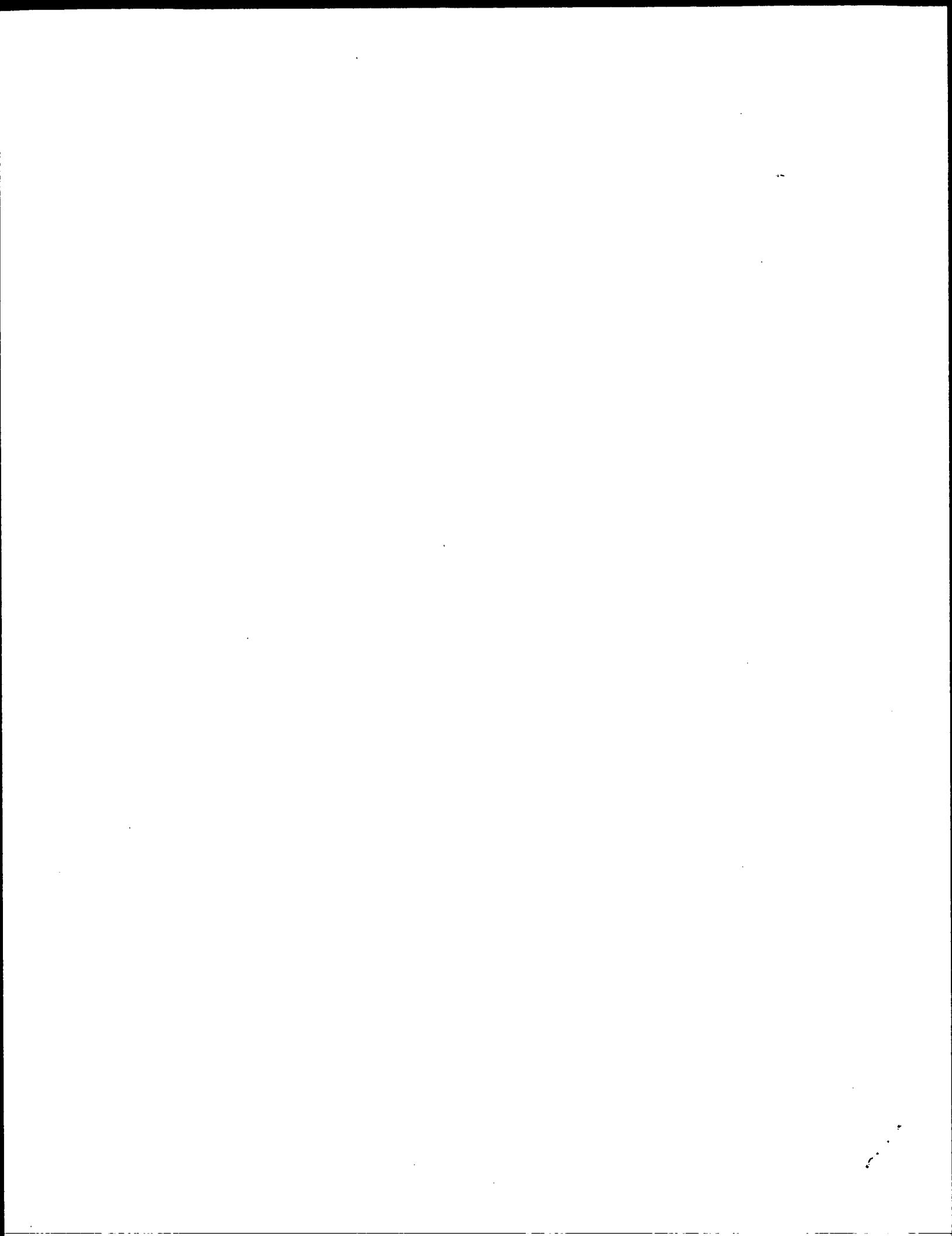
## RESULT 15

FAR3\_CALVO  
ID FAR3\_CALVO STANDARD; PRT; 9 AA.  
AC P41858;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE CallifMRamide 3.  
OS Calliphora vomitoria (Blue blowfly).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
OC Muscomorpha; Oestroidea; Calliphoridae; Calliphora.  
OX NCBI\_TaxID=27454;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Thoracic ganglion;  
RX MEDLINE=92196111; PubMed=1549595;  
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,  
RA Rehfeld J.F., Thorpe A.;  
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2  
RT neuropeptides (designated callifMRFamides) from the blowfly  
RT Calliphora vomitoria.";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).  
CC -!- FUNCTION: ABLE TO INDUCE FLUID SECRETION FROM THE ISOLATED  
CC SALIVARY GLAND OF CALLIPHORA.  
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
CC FAMILY.  
KW PIR; C41978; C41978.  
DR Neuropeptide; Amidation.  
FT MOD\_RES 9 9  
SQ SEQUENCE 9 AA; 1114 MW; 2F0B0699CAB6C5A7 CRC64;

Query Match 21.2%; Score 14; DB 1; Length 9;  
Best Local Similarity 44.4%; Pred. No. 1.1e+05;  
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 SLSFXFTKF 11  
| | | | |  
Db 1 SPSQDFMR 9

Search completed: February 26, 2003, 14:56:39  
Job time : 10 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:54:20 ; Search time 28 Seconds  
(without alignments)  
103.024 Million cell updates/sec

Title: US-09-476-485A-31  
Perfect score: 66  
Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 2683

Minimum DB seq length: 0  
Maximum DB seq length: 14

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | % Match | Query Length | DB ID     | Description         |
|------------|-------|---------|--------------|-----------|---------------------|
| 1          | 29    | 43.9    | 8            | 3 P87225  | P87225 saccharomyc  |
| 2          | 27    | 40.9    | 14           | 4 Q96KF9  | Q96KF9 homo sapien  |
| 3          | 24    | 36.4    | 9            | 4 Q14277  | Q14277 homo sapien  |
| 4          | 19    | 28.8    | 13           | 6 Q9TU76  | Q9TU76 ovis aries   |
| 5          | 19    | 28.8    | 14           | 2 Q9R518  | Q9R518 vibrio algi  |
| 6          | 18    | 27.3    | 9            | 2 Q8RKC6  | Q8RKC6 erwinia chr  |
| 7          | 18    | 27.3    | 14           | 15 Q9PYL3 | Q9PYL3 human t-cel  |
| 8          | 17    | 25.8    | 10           | 11 Q8VHM9 | Q8VHM9 mus musculus |
| 9          | 17    | 25.8    | 10           | 13 Q73588 | Q73588 gallus gall  |
| 10         | 17    | 25.8    | 12           | 4 Q96PK0  | Q96PK0 homo sapien  |
| 11         | 17    | 25.8    | 13           | 4 Q9UCT1  | Q9UCT1 homo sapien  |
| 12         | 17    | 25.8    | 13           | 11 Q63047 | Q63047 rattus norv  |
| 13         | 17    | 25.8    | 14           | 2 Q9LCS1  | Q9LCS1 bacillus su  |
| 14         | 17    | 25.8    | 14           | 2 Q8VQ14  | Q8VQ14 micrococcu   |
| 15         | 17    | 25.8    | 14           | 4 Q96Q62  | Q96Q62 homo sapien  |
| 16         | 17    | 25.8    | 14           | 10 P82435 | P82435 nicotiana t  |

|    |    |      |    |    |        |                    |
|----|----|------|----|----|--------|--------------------|
| 17 | 17 | 25.8 | 14 | 11 | Q9Z1H4 | Q9Z1H4 mus musculu |
| 18 | 17 | 25.8 | 14 | 12 | Q67112 | Q67112 influenzavi |
| 19 | 16 | 24.2 | 7  | 2  | Q07354 | Q07354 synechococc |
| 20 | 16 | 24.2 | 9  | 8  | Q8WGE6 | Q8WGE6 procambaru  |
| 21 | 16 | 24.2 | 9  | 13 | P83059 | P83059 bombina ori |
| 22 | 16 | 24.2 | 9  | 13 | P83058 | P83058 bombina var |
| 23 | 16 | 24.2 | 9  | 13 | P83057 | P83057 bombina var |
| 24 | 16 | 24.2 | 9  | 13 | P83056 | P83056 bombina var |
| 25 | 16 | 24.2 | 10 | 4  | O60912 | O60912 homo sapien |
| 26 | 16 | 24.2 | 10 | 6  | Q9TRC1 | Q9TRC1 bos taurus  |
| 27 | 16 | 24.2 | 10 | 10 | Q9S936 | Q9S936 beta vulgar |
| 28 | 16 | 24.2 | 11 | 2  | Q9AIZ8 | Q9AIZ8 carsonella  |
| 29 | 16 | 24.2 | 11 | 2  | Q9AIV6 | Q9AIV6 carsonella  |
| 30 | 16 | 24.2 | 11 | 7  | O78119 | O78119 oreochromis |
| 31 | 16 | 24.2 | 12 | 11 | Q9Z5V7 | Q9Z5V7 mus musculu |
| 32 | 16 | 24.2 | 13 | 2  | O82835 | O82835 synechococc |
| 33 | 16 | 24.2 | 14 | 2  | Q9R8L2 | Q9R8L2 chlamydia t |
| 34 | 16 | 24.2 | 14 | 2  | Q9R8J2 | Q9R8J2 chlamydia t |
| 35 | 16 | 24.2 | 14 | 2  | Q9R8J0 | Q9R8J0 chlamydia t |
| 36 | 16 | 24.2 | 14 | 2  | Q9R8I8 | Q9R8I8 chlamydia t |
| 37 | 16 | 24.2 | 14 | 2  | Q9R8I6 | Q9R8I6 chlamydia t |
| 38 | 16 | 24.2 | 14 | 2  | Q9R8I4 | Q9R8I4 chlamydia t |
| 39 | 16 | 24.2 | 14 | 2  | Q9R8I2 | Q9R8I2 chlamydia t |
| 40 | 16 | 24.2 | 14 | 2  | Q9R8I0 | Q9R8I0 chlamydia t |
| 41 | 16 | 24.2 | 14 | 2  | Q9R8H8 | Q9R8H8 chlamydia t |
| 42 | 16 | 24.2 | 14 | 2  | Q9R8H6 | Q9R8H6 chlamydia t |
| 43 | 16 | 24.2 | 14 | 2  | Q9R8H4 | Q9R8H4 chlamydia t |
| 44 | 16 | 24.2 | 14 | 2  | Q9R8H2 | Q9R8H2 chlamydia t |
| 45 | 16 | 24.2 | 14 | 2  | Q9R8H0 | Q9R8H0 chlamydia t |

## ALIGNMENTS

### RESULT 1

P87225  
ID P87225 PRELIMINARY; PRT; 8 AA.  
AC P87225;  
DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE GIN11 protein (Fragment).  
GN GIN11 OR YLL065W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wedler H., Wedler E., Scharfe M., Wambutt R.;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MIPS;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; 273159; CAA97518.2; -.  
DR SGD; S0003988; GIN11.  
FT NON\_TER 1  
SQ SEQUENCE 8 AA; 1019 MW; 4E21A9C449D5B73B CRC64;

Query Match 43.9%; Score 29; DB 3; Length 8;  
Best Local Similarity 85.7%; Pred. No. 6.7e+05;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 LSFXTFK 10  
||| |||  
DB 2 LSFNFTK 8

### RESULT 2

Q96KF9  
ID Q96KF9 PRELIMINARY; PRT; 14 AA.  
AC Q96KF9;

```

DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DE Putative magnesium transporter (Fragment).
GN MRS2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BL00D;
RX MEDLINE=21295035; PubMed=11401429;
RA Zsuzka G., Gregan J., Schweyen R.J.;
RT "The human mitochondrial MRS2 protein functionally substitutes for its
RL yeast homologue, a candidate magnesium transporter.";
DR Genomics 72:158-168(2001).
DR EMBL: AF293077; AAK38616.1; -.
FT NON_TER 1 14
FT NON_TER 14 14
SQ SEQUENCE 14 AA; 1537 MW; C6344FF1E984AFB2 CRC64;

Query Match 40.9%; Score 27; DB 4; Length 14;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 FXTKFD 12
Db 5 FVTKFD 11

RESULT 3
Q14277 PRELIMINARY; PRT; 9 AA.
AC Q14277;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DE RET protein short form (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94071887; PubMed=7902707;
RA Ceccherini I., Bocciardi R., Luo Y., Pasini B., Hofstra R.,
RA Takahashi M., Romeo G.;
RT "Exon structure and flanking intronic sequences of the human RET
RT proto-oncogene.";
RL Biochem. Biophys. Res. Commun. 196:1288-1295(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94366753; PubMed=8084609;
RA Ceccherini I., Hofstra R., Yin L., Stulp R., Barone V., Stelwagen T.,
RA Bocciardi R., Nijveen H., Bolino A., Seri M., Ronchetto P., Pasini B.,
RA Bozzano M., Buys C., Romeo G.;
RT "DNA polymorphisms and conditions for SSCP analysis of the 20 exons of
RT the ret proto-oncogene.";
RL Oncogene 9:3025-3029(1994).
DR EMBL: U11532; AAC50102.1; -.
FT NON_TER 1 1
FT NON_TER 1 1
SQ SEQUENCE 9 AA; 1134 MW; 034E11A9CDD1F5B0 CRC64;

Query Match 36.4%; Score 24; DB 4; Length 9;
Best Local Similarity 50.0%; Pred. No. 6.7e+05;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 LSFYTKF 11
Db 2 ISHAFTRF 9

```

```

RESULT 4
Q9TU76 PRELIMINARY; PRT; 13 AA.
ID Q9TU76;
AC Q9TU76;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE Interleukin 1 alpha (Fragment).
GN IL1A.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99397042; PubMed=10467711;
RA Maddox J.F., Hawken R.J., Matthew P., Davies K.P.;
RT "Single strand conformational polymorphisms (SSCPs) in the ovine IL1A
RT and IL6 genes.";
RL Anim. Genet. 30:317-318(1999).
DR EMBL: AF117652; AAD25050.1; -.
FT NON_TER 1 1
FT NON_TER 13 13
SQ SEQUENCE 13 AA; 1520 MW; C26BC6198305DB5D CRC64;

Query Match 28.8%; Score 19; DB 6; Length 13;
Best Local Similarity 75.0%; Pred. No. 5e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KFDL 13
Db 1 KFDL 4

RESULT 5
Q9R518 PRELIMINARY; PRT; 14 AA.
ID Q9R518;
AC Q9R518;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
DE L-2,4-diaminobutyrate decarboxylase (Fragment).
OS Vibrio alginolyticus.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=663;
RN [1]
RP SEQUENCE.
RX MEDLINE=92381494; PubMed=1512577;
RA Yamanoto S., Tsuzaki Y., Tougou K., Shinoda S.;
RT "Purification and characterization of L-2,4-diaminobutyrate
RT decarboxylase from Acinetobacter calcoaceticus.";
RL J. Gen. Microbiol. 138:1461-1465(1992).
SQ SEQUENCE 14 AA; 1643 MW; 9F1B13DD35168ABA CRC64;

Query Match 28.8%; Score 19; DB 2; Length 14;
Best Local Similarity 50.0%; Pred. No. 5.4e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 9 TKFDLD 14
Db 2 TAFEVD 7

RESULT 6
Q8RK6 PRELIMINARY; PRT; 9 AA.
ID Q8RK6;
AC Q8RK6;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Expr protein (Fragment).
GN EXPR.

```

OS Erwinia chrysanthemi.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Pectobacterium.  
 OX NCBI\_TaxID=556;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=3937;  
 RA Reverchon S.;  
 RT "Identification of a lysA-like gene required for virulence factors  
 synthesis in Erwinia chrysanthemi.";  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ438189; CAD27339.1; -  
 FT NON\_TER 9  
 SQ SEQUENCE 9 AA; 999 MW; 9A8BC455B9D5B045 CRC64;  
 Query Match 27.3%; Score 18; DB 2; Length 9;  
 Best Local Similarity 33.3%; Pred. No. 6.7e+05;  
 Matches 3; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 4 LSFXTKFD 12  
 : | | : |  
 Db 1 MSISFSNVD 9  
 RESULT 7  
 Q9PYL3 PRELIMINARY; PRT; 14 AA.  
 AC Q9PYL3;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
 DE Transcriptional transactivator Tax (Fragment).  
 GN TAX.  
 OS Human T-cell leukemia virus type II (HTLV-II).  
 OC Viruses; Retroviridae; Retroviridae; Deltaretrovirus.  
 OX NCBI\_TaxID=11909;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H2BC1;  
 RX MEDLINE=20014165; PubMed=10548129;  
 RA Peters A.A., Oger J.J., Coulthart M.B., Waters D.J., Cummings H.J.,  
 RA Dekaban G.A.;  
 RT "An apparent case of human T-cell lymphotropic virus type II (HTLV-II)-  
 associated neurological disease: a clinical, molecular, and  
 phylogenetic characterisation.";  
 RL J. Clin. Virol. 14:37-50(1999).  
 DR EMBL; AF115495; AAF15550.1; -  
 FT NON\_TER 1  
 SQ SEQUENCE 14 AA; 1567 MW; 7353D4CA8CC6041B CRC64;  
 Query Match 27.3%; Score 18; DB 15; Length 14;  
 Best Local Similarity 40.0%; Pred. No. 8.5e+03;  
 Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 QY 5 SFXFTKFDLD 14  
 : | | : |  
 Db 1 SILENKBEAD 10  
 RESULT 8  
 Q8VHM9 PRELIMINARY; PRT; 10 AA.  
 AC Q8VHM9;  
 DT 01-MAR-2002 (TReMBLrel. 20, Created)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
 DE Interferon receptor 2a (Fragment).  
 GN IFNAR2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6;  
 RA Hardy M.P., Hertzog P.J., Owczarek C.M.;  
 RT "The genomic structure and expression patterns of the gene encoding  
 the second chain of the murine interleukin 10 receptor, IL-10R2.";  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6;  
 RA Hardy M.P., Hertzog P.J., Owczarek C.M.;  
 RT "The organization, transcriptional regulation and chromosomal  
 localization of the locus encoding the gene for the murine type I  
 interferon receptor, Ifnar2.";  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF440786; AAL40944.1; -  
 KW Receptor.  
 FT NON\_TER 1  
 SQ SEQUENCE 10 AA; 1139 MW; 322A36A72449C737 CRC64;  
 Query Match 25.8%; Score 17; DB 11; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 9.5e+03;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 11 FDL 14  
 : | | : |  
 Db 5 FNLD 8  
 RESULT 9  
 O73588 PRELIMINARY; PRT; 10 AA.  
 ID O73588;  
 AC O73588;  
 DT 01-AUG-1998 (TReMBLrel. 07, Created)  
 DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE Engrailed-3 (Fragment).  
 GN EN-3.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WHITE LEGHORN; TISSUE=EMBRYO;  
 RX MEDLINE=98141813; PubMed=9473273;  
 RA Peale F.V. Jr., Mason K., Hunter A.W., Bothwell M.;  
 RT "Multiplex display polymerase chain reaction amplifies and resolves  
 related sequences sharing a single moderately conserved domain.";  
 RL Anal. Biochem. 256:158-168(1998).  
 DR EMBL; U26148; AAC06186.1; -  
 FT NON\_TER 1  
 SQ SEQUENCE 10 AA; 1118 MW; 73C0BE14735B72B CRC64;  
 Query Match 25.8%; Score 17; DB 13; Length 10;  
 Best Local Similarity 80.0%; Pred. No. 9.5e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AQSLS 5  
 : | | : |  
 Db 1 AQELS 5  
 RESULT 10  
 Q96PK0 PRELIMINARY; PRT; 12 AA.  
 ID Q96PK0;  
 AC Q96PK0;  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE COPG2 (Fragment).  
 GN COPG2.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bonora E., Bacchelli E., Levy E.R., Blasi F., Marlow A., Monaco A.P.,  
 RA Maestrini E.;  
 RT "Mutation screening and imprinting analysis of four candidate genes  
 RT for autism in the 7q32 region.";  
 RL Mol. Psychiatry 0:0-0(2001).  
 DR EMBL: AF324497; AAL1778.1; -.  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1498 MW; C344E6689A333AA9 CRC64;  
 Query Match 25.8%; Score 17; DB 4; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+04;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 10 KFD 12  
 Db 4 KFD 6  
 RESULT 11  
 ID Q9UCT1 PRELIMINARY; PRT; 13 AA.  
 AC Q9UCT1;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Platelet glycoprotein GPIIC (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92061944; PubMed=1953640;  
 RA Catimel B., Parmentier S., Leung L.L., McGregor J.L.;  
 RT "Separation of important new platelet glycoproteins (GPIa, GPIIb,  
 RT GPIc, GPIIb and GPIIb-100) by f.p.l.c. Characterization by monoclonal  
 RT antibodies and gas-phase sequencing.";  
 RL Biochem. J. 279:419-425(1991).  
 FT NON\_TER 1 1  
 FT NON\_TER 13 13  
 SQ SEQUENCE 13 AA; 1620 MW; 79740A62649AB04 CRC64;  
 Query Match 25.8%; Score 17; DB 4; Length 13;  
 Best Local Similarity 75.0%; Pred. No. 1.2e+04;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 11 FDL 14  
 Db 1 FDL 4  
 RESULT 12  
 ID Q63047 PRELIMINARY; PRT; 13 AA.  
 AC Q63047;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE AMP deaminase (Fragment).  
 GN AMPD1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=SOLEUS MUSCLE;

RX MEDLINE=90377216; PubMed=2398891;  
 RA Mineo I., Clarke P.R.H., Sabina R.L., Holmes E.W.;  
 RT "A novel pathway for alternative splicing: Identification of an RNA  
 RT intermediate that generates an alternative 5' splice donor site not  
 RT present in the primary transcript of AMPD1.";  
 RL Mol. Cell. Biol. 10:5271-5278(1990).  
 DR EMBL: M38689; AAA40727.1; -.  
 FT NON\_TER 13 13  
 SQ SEQUENCE 13 AA; 1524 MW; 526C5A93EF6201A7 CRC64;  
 Query Match 25.8%; Score 17; DB 11; Length 13;  
 Best Local Similarity 42.9%; Pred. No. 1.2e+04;  
 Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 Qy 6 FXFTKFD 12  
 Db 4 FKLTEID 10  
 RESULT 13  
 ID Q9LCSI PRELIMINARY; PRT; 14 AA.  
 AC Q9LCSI;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE Levan sucrose (Fragment).  
 GN SACB.  
 OS Bacillus subtilis.  
 OG Plasmid p1257.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
 OC Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Maldonado R., Casadesus J.;  
 RT "Identification of IS210 in Azotobacter vinelandii: a novel,  
 RT functional insertion element member of the IS5 family.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ249381; CAB76429.1; -.  
 KW Plasmid.  
 FT NON\_TER 1 1  
 FT NON\_TER 14 14  
 SQ SEQUENCE 14 AA; 1623 MW; 8DC9108BA1B18745 CRC64;  
 Query Match 25.8%; Score 17; DB 2; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+04;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 10 KFD 12  
 Db 1 KFD 3  
 RESULT 14  
 ID Q8VQ14 PRELIMINARY; PRT; 14 AA.  
 AC Q8VQ14;  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE ERMML leader peptide.  
 OS Micrococcus luteus (Micrococcus lysodeikticus).  
 OG Plasmid pMEC2.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Micrococcales; Micrococcaceae; Micrococcus.  
 OX NCBI\_TaxID=1270;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAW843;  
 RA Liebl W., Kloos W.E., Ludwig W.;  
 RT "Plasmid-borne macrolide-lincosamide-streptogramin B (MLS) resistance  
 RT in Micrococcus luteus.";



RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF462611; AAL68826.1; -  
 KW Plasmid.

SQ SEQUENCE 14 AA; 1625 MW; C6BF8E5F5CD58BCD CRC64;

Query Match 25.8%; Score 17; DB 2; Length 14;  
 Best Local Similarity 44.4%; Pred. No. 1.3e+04;  
 Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 SLSFXFTKF 11

1 1 1 1 1 1

Db 3 SPSTAVTRF 11

# RESULT 15

ID Q96Q62

AC Q96Q62 PRELIMINARY; PRT; 14 AA.

DT 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE Mitochondrial ribosomal protein S11 (Fragment).

GN MRPS11.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

EX MEDLINE=21429115; PubMed=11543634;

RA Kenmochi N., Suzuki T., Uechi T., Magoori M., Kuniba M., Higa S.,

RA Watanabe K., Tanaka T.;

RT "The human mitochondrial ribosomal protein genes: Mapping of 54 genes

to the chromosomes and implications for human disorders.";

RL Genomics 77:65-70(2001).

DR EMBL; AB051349; BAB54939.2; -.

KW Ribosomal protein.

FT NON\_TER 1

FT NON\_TER 14

SQ SEQUENCE 14 AA; 1585 MW; C07121F2234438DD CRC64;

Query Match

Best Local Similarity 25.8%; Score 17; DB 4; Length 14;

Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 QSLFXFTKF 11

1 1 1 1 1 1

Db 5 QNAAPSHTKF 14

Search completed: February 26, 2003, 14:57:15

Job time : 30 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:51:20 ; Search time 34 seconds  
(without alignments)  
54.868 Million cell updates/sec

Title: US-09-476-485A-31  
Perfect score: 66  
Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 228063

Minimum DB seq length: 0  
Maximum DB seq length: 14

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | ID | Description        |
|------------|-------|-------------|----|--------------------|
| 1          | 65    | 98.5        | 14 | AAW61498           |
| 2          | 64    | 97.0        | 14 | AAW67818           |
| 3          | 64    | 97.0        | 14 | AAW62895           |
| 4          | 55    | 83.3        | 12 | AAW61497           |
| 5          | 47    | 71.2        | 14 | AAW62896           |
| 6          | 25    | 37.9        | 8  | AAW62906           |
| 7          | 25    | 37.9        | 8  | AAW63026           |
| 8          | 25    | 37.9        | 9  | AAW46829           |
| 9          | 25    | 37.9        | 9  | AAW62907           |
| 10         | 25    | 37.9        | 9  | AAW63027           |
|            |       |             |    | Pylartin protein,  |
|            |       |             |    | Flk2 ligand N-term |
|            |       |             |    | Antigenic peptide  |
|            |       |             |    | Pylartin protein,  |
|            |       |             |    | Peptide derived fr |
|            |       |             |    | PB-cadherin cell a |
|            |       |             |    | PB-cadherin cell a |
|            |       |             |    | Immunogenic peptid |
|            |       |             |    | PB-cadherin cell a |
|            |       |             |    | PB-cadherin cell a |

|    |    |      |    |    |          |                     |
|----|----|------|----|----|----------|---------------------|
| 11 | 25 | 37.9 | 9  | 22 | AAW61498 | Human TADG-15 pept  |
| 12 | 25 | 37.9 | 9  | 22 | AAW61498 | Human TADG-15 pept  |
| 13 | 25 | 37.9 | 10 | 21 | AAW78322 | Anti-zeta-chain an  |
| 14 | 25 | 37.9 | 10 | 21 | AAW62908 | PB-cadherin cell a  |
| 15 | 25 | 37.9 | 10 | 21 | AAW63028 | PB-cadherin cell a  |
| 16 | 25 | 37.9 | 11 | 21 | AAW12876 | Protein kinase pep  |
| 17 | 25 | 37.9 | 12 | 18 | AAW27590 | Anti-TNF-alpha ant  |
| 18 | 25 | 37.9 | 12 | 18 | AAW39757 | Anti-IL12 antibody  |
| 19 | 25 | 37.9 | 12 | 22 | AAU05005 | N-terminal peptide  |
| 20 | 25 | 37.9 | 13 | 23 | AAU86065 | Human glucocerebro  |
| 21 | 24 | 36.4 | 7  | 16 | AAW1396  | VPL peptide used t  |
| 22 | 24 | 36.4 | 7  | 19 | AAW38291 | Poliovirus amino a  |
| 23 | 24 | 36.4 | 7  | 22 | AAW37816 | Peptide unique to   |
| 24 | 24 | 36.4 | 9  | 21 | AAW39807 | Anti-hil12 antibod  |
| 25 | 24 | 36.4 | 9  | 22 | AAE10843 | Cytomegalovirus pp  |
| 26 | 24 | 36.4 | 10 | 15 | AAW49439 | Immunomodulatory p  |
| 27 | 24 | 36.4 | 10 | 19 | AAW76007 | LM609 grafted anti  |
| 28 | 24 | 36.4 | 10 | 22 | ABB56106 | Vascular dementia-  |
| 29 | 24 | 36.4 | 10 | 22 | AAU28788 | DPI tryptic digest  |
| 30 | 24 | 36.4 | 10 | 22 | AAU26436 | Depression-Associat |
| 31 | 24 | 36.4 | 10 | 22 | AAW61365 | LM609 VH CDRI1 pept |
| 32 | 24 | 36.4 | 10 | 23 | ABG67865 | Human ADPI tryptic  |
| 33 | 24 | 36.4 | 12 | 11 | AAW06818 | Cell surface recep  |
| 34 | 24 | 36.4 | 12 | 13 | AAW20117 | MHC Class I-derive  |
| 35 | 24 | 36.4 | 12 | 15 | AAW49477 | HLA-A2 position 29  |
| 36 | 24 | 36.4 | 12 | 15 | AAW49390 | HLA-A2 position 31  |
| 37 | 24 | 36.4 | 12 | 15 | AAW49401 | HLA-Cw9 position 3  |
| 38 | 24 | 36.4 | 12 | 16 | AAW69608 | MHC class I-derive  |
| 39 | 24 | 36.4 | 12 | 20 | AAW25343 | Human p60 tryptic   |
| 40 | 24 | 36.4 | 13 | 11 | AAW06819 | Cell surface recep  |
| 41 | 24 | 36.4 | 13 | 13 | AAW20118 | MHC Class I-derive  |
| 42 | 24 | 36.4 | 13 | 16 | AAW69609 | MHC class I-derive  |
| 43 | 24 | 36.4 | 13 | 19 | AAW77284 | P. damsela beta-ga  |
| 44 | 24 | 36.4 | 13 | 19 | AAW44131 | HLA-A2 peptide (aa  |
| 45 | 24 | 36.4 | 14 | 15 | AAW49400 | HLA-Cw9 position 3  |

ALIGNMENTS

RESULT 1  
AAW61498  
ID AAW61498 standard; peptide; 14 AA.  
AC AAW61498;  
XX  
XX  
19-OCT-1998 (first entry)  
DT  
XX  
XX  
Pylartin protein, peptide chain beta (ii).  
DE  
XX  
XX  
Pylartin protein; progenitor cell; haematopoietic system; cancer;  
KW engraftation; haematologic disease; sickle cell anaemia; thalassemia.  
XX  
XX  
Leguminosae.  
XX  
XX  
WO9825457-A1.  
XX  
XX  
18-JUN-1998.  
XX  
XX  
09-DEC-1997; 97WO-US22486.  
XX  
XX  
28-MAR-1997; 97US-0825369.  
XX  
XX  
09-DEC-1996; 96US-0762537.  
XX  
XX  
(IMCL-) IMCLONE SYSTEMS INC.  
XX  
XX  
Moore JG;  
XX  
XX  
WPI; 1998-348161/30.  
XX  
XX  
Pylartin proteins which preserve progenitor cells - useful for  
PT haematopoietic therapies in cancer treatment or for treating  
PT haematologic diseases

XX Claim 4; Page 28; 46pp; English.

XX The peptides AAW61497-W61502 can be used to form pylartin proteins which

CC preserve progenitor cells (pc). The proteins can be used for protecting

CC the integrity of the haematopoietic processes in vivo and as adjuncts in

CC therapeutic treatments related to cancer and other diseases which can

CC otherwise adversely impact upon the haematopoietic system. Since the

CC proteins bind specifically to primitive PCs, they can also be used for

CC the identification and localisation of PCs. The methods can be used for

CC e.g. expanding PC populations ex vivo to increase chances of

CC engraftation, improving conditions for transporting and storing PCs and

CC for removing a fundamental barrier thereby enabling gene therapy to

CC treat and cure a broad range of life-threatening haematologic diseases

CC such as sickle cell anaemia and thalassemia.

XX

SQ Sequence 14 AA;

Query Match 98.5%; Score 65; DB 19; Length 14;

Best Local Similarity 92.9%; Pred. No. 9.1e-06;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14

DB 1 AQSLSFSFTKFDLD 14

RESULT 2

AAR67818

ID AAR67818 standard; peptide; 14 AA.

XX

AC AAR67818;

XX

DT 18-AUG-1995 (first entry)

DE Flk2 ligand N-terminal sequence.

XX

KW human Flk2 receptor protein-tyrosine-kinase ligand; peripheral

KW periperal blood leukocyte conditioned medium;

KW bone marrow disorder diagnosis; hematopoietic stem cell;

KW proliferation; differentiation.

XX

OS Homo sapiens.

XX

PH Key Location/Qualifiers

FT misc\_difference 7 /note= "any amino acid"

FT

XX WO9500554-A.

XX

PN 05-JAN-1995.

PD

XX

PF 17-JUN-1994; 94WO-US06944.

XX

PR 18-JUN-1993; 93US-0080244.

XX

PR 21-JUN-1993; 93US-0081508.

XX

PR 23-NOV-1993; 93US-0157490.

XX

PA (UYPR-) UNIV PRINCETON.

XX

XX Lemischka IR;

XX

XX WPI; 1995-052014/07.

XX

PT Ligand for receptor protein tyrosine kinase - useful for the

PT stimulation of primitive haematopoietic stem cells causing

PT proliferation and/or differentiation

XX

PS Claim 1; Page 105; 131pp; English.

XX

CC The sequence corresponds to the N-terminal region of a human Flk2

CC receptor protein-tyrosine-kinase ligand, isolated from

CC phytohemagglutinin-stimulated human peripheral blood leukocyte

CC

CC tissue culture conditioned medium. The ligand may be used in

CC diagnosis of bone marrow disorders, and to stimulate the

CC proliferation and/or differentiation of primitive hematopoietic stem

CC cells. The ligand binds to a receptor protein-tyrosine-kinase

CC expressed in primitive but not mature mammalian hematopoietic cells.

XX

SQ Sequence 14 AA;

Query Match 97.0%; Score 64; DB 16; Length 14;

Best Local Similarity 100.0%; Pred. No. 1.4e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14

DB 1 AQSLSFXFTKFDLD 14

RESULT 3

AAG62895

ID AAG62895 standard; peptide; 14 AA.

XX

AC AAG62895;

XX

DT 17-SEP-2001 (first entry)

DE Antigenic peptide derived from a french bean FRIL polypeptide.

XX

KW FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;

KW progenitor cell preservation factor; radiotherapy; chemotherapy;

KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;

KW severe combined immunodeficiency; aplastic anemia; tissue repair.

XX

OS Phaseolus vulgaris.

XX

PH Key Location/Qualifiers

FT Misc-difference 7 /label= Asn, Cys, Ser

FT

XX WO200149851-A1.

XX

PD 12-JUL-2001.

XX

PF 30-DEC-1999; 99WO-US31307.

XX

PR 30-DEC-1999; 99WO-US31307.

XX

PA (PHYL-) PHYLOGIX LLC.

XX

PI Colucci MG, Chrispeels MJ, Moore JG;

XX

DR WPI; 2001-441882/47.

XX

PT Legume Progenitor cell preservation factors for in vivo or ex vivo

PT preservation of hematopoietic progenitor cells and as therapeutics for

PT alleviating/reducing progenitor cell-depleting activity of cancer

XX therapeutics -

XX

PS Example 5; Page 72; 173pp; English.

XX

CC The present sequence is derived from FRIL (Flk2/Flt3 tyrosine kinase

CC receptor-interacting lectin), and is used to raise antibodies. The

CC specification describes a composition of one or more members

CC of FRIL family of progenitor cell preservation factors. The composition

CC is useful for alleviating or reducing the hematopoietic progenitor

CC cell-depleting activity of a therapeutic treatment, including

CC radiotherapeutic and/or chemotherapeutic treatments. Administration of

CC FRIL compositions to a patient prior to treatment of the patient with

CC a therapeutic treatment having a hematopoietic progenitor cell-depleting

CC activity alleviates or reduces the hematopoietic progenitor

CC cell-depleting activity of the therapeutic treatment in the patient.

CC FRIL family members are useful for isolating population of progenitor

CC cells, hemangioblasts, and mesenchymal stem cells. The composition is

CC administered to reduce progenitor cell depleting effects of

chemotherapeutics, so that the patient can receive a higher dose of the chemotherapeutic and preferably recover from cancer. It is also administered to patients having, or predisposed to developing a condition where the patients hematopoietic progenitor cells are depleted, such as severe combined immunodeficiency or aplastic anemia. The isolated mesenchymal cells are useful for tissue repair.

XX SQ Sequence 14 AA;

Query Match 97.0%; Score 64; DB 22; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AQSLSFXFTKFDLD 14  
|||||

Db 1 AQSLSFXFTKFDLD 14

# RESULT 4

AAW61497  
ID AAW61497 standard; peptide; 12 AA.

XX AAW61497;

DT 19-OCT-1998 (first entry)

DE Pylartin protein, peptide chain beta (i).

XX Pylartin protein; progenitor cell; hematopoietic system; cancer;  
KW engraftation; hematologic disease; sickle cell anaemia; thalassemia.

XX Leguminosae.

XX W09825457-A1.

XX 18-JUN-1998.

XX 09-DEC-1997; 97WO-US22486.

XX 28-MAR-1997; 97US-0825369.

XX 09-DEC-1996; 96US-0762537.

XX (IMCL-) IMCLONE SYSTEMS INC.

XX Moore JG;

XX WPI; 1998-348161/30.

XX Pylartin proteins which preserve progenitor cells - useful for  
PT hematopoietic therapies in cancer treatment or for treating  
PT hematologic diseases

XX Claim 1; Page 28; 46pp; English.

XX The peptides AAW61497-W61502 can be used to form pylartin proteins which  
CC preserve progenitor cells (pc). The proteins can be used for protecting  
CC the integrity of the hematopoietic processes in vivo and as adjuncts in  
CC therapeutic treatments related to cancer and other diseases which can  
CC otherwise adversely impact upon the hematopoietic system. Since the  
CC proteins bind specifically to primitive PCs, they can also be used for  
CC the identification and localization of PCs. The methods can be used for  
CC e.g. expanding PC populations ex vivo to increase chances of  
CC engraftation, improving conditions for transporting and storing PCs and  
CC for removing a fundamental barrier thereby enabling gene therapy to  
CC treat and cure a broad range of life-threatening hematologic diseases  
CC such as sickle cell anaemia and thalassemia.

XX SQ Sequence 12 AA;

Query Match 83.3%; Score 55; DB 19; Length 12;  
Best Local Similarity 91.7%; Pred. No. 0.00061;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AQSLSFXFTKFD 12  
|||||

Db 1 AQSLSFXFTKFD 12

# RESULT 5

AAG62896  
ID AAG62896 standard; peptide; 14 AA.

XX AAG62896;

XX 17-SEP-2001 (first entry)

DE Peptide derived from a french bean FRIL polypeptide.

XX FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;  
KW progenitor cell preservation factor; radiotherapy; chemotherapy;  
KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;  
KW severe combined immunodeficiency; aplastic anemia; tissue repair.

XX Phaseolus vulgaris.

XX Key Location/Qualifiers  
FH Misc-difference 7  
FT /note= "not known"

XX W0200149851-A1.

XX 12-JUL-2001.

XX 30-DEC-1999; 99WO-US31307.

XX 30-DEC-1999; 99WO-US31307.

XX (PHYL-) PHYLOGIX LLC.

XX Colucci MG, Chrispeels MJ, Moore JG;

XX WPI; 2001-441882/47.

XX Legume progenitor cell preservation factors for in vivo or ex vivo  
PT preservation of hematopoietic progenitor cells and as therapeutics for  
PT alleviating/reducing progenitor cell-depleting activity of cancer  
PT therapeutics

XX Example 5; Page 75; 173pp; English.

XX The present sequence is derived from a FRIL (Flk2/Flt3 tyrosine kinase  
CC receptor-interacting lectin) polypeptide. The specification describes a  
CC composition of one or more members of FRIL family of progenitor cell  
CC preservation factors. The composition is useful for alleviating or  
CC reducing the hematopoietic progenitor cell-depleting activity of a  
CC therapeutic treatment, including radiotherapeutic and/or  
CC chemotherapeutic treatments. Administration of FRIL compositions to a  
CC patient prior to treatment of the patient with a therapeutic treatment  
CC having a hematopoietic progenitor cell-depleting activity alleviates or  
CC reduces the hematopoietic progenitor cell-depleting activity of the  
CC therapeutic treatment in the patient. FRIL family members are useful for  
CC isolating population of progenitor cells, hemangioblasts, and mesenchymal  
CC stem cells. The composition is administered to reduce progenitor cell  
CC depleting effects of chemotherapeutics, so that the patient can receive  
CC a higher dose of the chemotherapeutic and preferably recover from cancer.  
CC It is also administered to patients having, or predisposed to developing  
CC a condition where the patients hematopoietic progenitor cells are  
CC depleted, such as severe combined immunodeficiency or aplastic anemia.  
CC The isolated mesenchymal cells are useful for tissue repair.

XX SQ Sequence 14 AA;

Query Match 71.2%; Score 47; DB 22; Length 14;  
Best Local Similarity 85.7%; Pred. No. 0.024;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14  
 DB 1 AQSLSFXFTKDALD 14

RESULT 6  
 AAY62906  
 ID AAY62906 standard; Peptide; 8 AA.  
 XX AAY62906;  
 AC AAY62906;  
 XX DT 02-MAR-2000 (first entry)  
 XX PB-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:2462.

XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT Modified-site 1.8  
 FT /note= "the terminal residues are condensed with each  
 FT other to form a cyclic peptide"

PN WO9957149-A2.  
 XX  
 PD 11-NOV-1999.  
 XX  
 XX 05-MAY-1999; 99WO-CA00363.  
 XX  
 XX 05-MAY-1998; 98US-0073040.  
 PR 06-NOV-1998; 98US-0187859.  
 PR 20-JAN-1999; 99US-0234395.  
 PR 08-MAR-1999; 99US-0264516.  
 XX  
 PA (ADHE-) ADHEREX TECHNOLOGIES INC.  
 XX  
 XX Blaschuk OW, Gour BJ, Byers S;  
 PI WPI; 2000-038791/03.  
 XX  
 XX New cadherin modulating agents, used for modulating nonclassical  
 PT cadherin-mediated functions for treating e.g. cancers, obesity,  
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
 PT disease -  
 XX  
 XX Claim 72; Page 195; 252pp; English.

XX The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound  
 CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue

CC in a mammal. They can also be used for treating e.g. psoriasis,  
 CC arthritis, age-related macular degeneration, multiple sclerosis and  
 CC diabetes. The products can also be used for detection and diagnosis and  
 CC in bioreactors. AAY60592 to AAY64572 represent specifically claimed  
 CC peptides, and AAY64573 to AAY64643 and AAY33183 to AAY33186 represent  
 CC sequences used in the exemplification of the present invention.  
 XX  
 XX Sequence 8 AA;  
 XX Query Match 37.9%; Score 25; DB 21; Length 8;  
 XX Best Local Similarity 80.0%; Pred. No. 7.8e+05;  
 XX Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 KFDLD 14  
 DB 1 KFDID 5  
 |||:|

RESULT 7  
 AAY63026  
 ID AAY63026 standard; Peptide; 8 AA.  
 XX AAY63026;  
 AC AAY63026;  
 XX DT 02-MAR-2000 (first entry)  
 XX PB-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:2582.

XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT Modified-site 1.8  
 FT /note= "the terminal residues are condensed with each  
 FT other to form a cyclic peptide"

PN WO9957149-A2.  
 XX  
 PD 11-NOV-1999.  
 XX  
 XX 05-MAY-1999; 99WO-CA00363.  
 XX  
 XX 05-MAY-1998; 98US-0073040.  
 PR 06-NOV-1998; 98US-0187859.  
 PR 20-JAN-1999; 99US-0234395.  
 PR 08-MAR-1999; 99US-0264516.  
 XX  
 PA (ADHE-) ADHEREX TECHNOLOGIES INC.  
 XX  
 XX Blaschuk OW, Gour BJ, Byers S;  
 PI WPI; 2000-038791/03.  
 XX  
 XX New cadherin modulating agents, used for modulating nonclassical  
 PT cadherin-mediated functions for treating e.g. cancers, obesity,  
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
 PT disease -  
 XX  
 XX Claim 72; Page 197; 252pp; English.

XX The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound  
 CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue

CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound  
 CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue  
 CC in a mammal. They can also be used for treating e.g. psoriasis,  
 CC arthritis, age-related macular degeneration, multiple sclerosis and  
 CC diabetes. The products can also be used for detection and diagnosis and  
 CC in bioreactors. AAY60592 to AAY64572 represent specifically claimed  
 CC peptides, and AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent  
 CC sequences used in the exemplification of the present invention.

XX Sequence 8 AA;

Query Match 37.9%; Score 25; DB 21; Length 8;  
 Best Local Similarity 80.0%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KFDLD 14  
 ||||  
 Db 1 KFDID 5

RESULT 8

AAAY46829  
 ID AAY46829 standard; Peptide; 9 AA.

XX AC AAY46829;

XX DT 01-DEC-1999 (first entry)

XX DE Immunogenic peptide having a human leukocyte antigen binding motif #1440.

XX KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;  
 KW immune response; T cell activation; major histocompatibility complex;  
 KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;  
 KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;  
 KW vaccine; immunisation.

XX OS Synthetic.  
 OS Homo sapiens.

XX PN WO5945954-A1.

XX PD 16-SEP-1999.

XX PF 13-MAR-1998; 98WO-US05039.

XX PR 13-MAR-1998; 98WO-US05039.

XX PA (EPIM-) EPIMUNE INC.

XX PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;

XX WPI; 1999-551214/46.

XX DR New immunogenic peptides with HLA binding motif, useful in treatment  
 XX and diagnosis of cancers and viral diseases -

XX PS Claim 1; Page 86; 150pp; English.

XX CC AAY45390 to AAY48214 represent specifically claimed immunogenic peptides  
 CC having a human major histocompatibility complex (MHC) Class I (also  
 CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic  
 CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes

CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell  
 CC response against the antigen from which the peptide is derived.  
 CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are  
 CC normally induced by an antigen in the form of a peptide fragment bound  
 CC to a HLA molecule, rather than the intact foreign antigen itself, and  
 CC are particularly important in tumour rejection and in fighting viral  
 CC infections. The peptides are therefore useful therapeutically to treat  
 CC or prevent viral infections and cancers in mammals (especially humans)  
 CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.  
 CC They can be administered as vaccines to elicit an immune response in  
 CC individuals susceptible or otherwise at risk of viral infection or  
 CC cancer, or used to treat chronic or acute conditions. They are also  
 CC useful diagnostically, and can be used to induce a cytotoxic T cell  
 CC response, by contacting a cytotoxic T cell with the peptide e.g. to  
 CC produce CTLs ex vivo for infusion back into a patient. The  
 CC polynucleotides encoding the immunogenic peptides are also useful  
 CC therapeutically and for immunisation as above.

XX Sequence 9 AA;

Query Match 37.9%; Score 25; DB 20; Length 9;  
 Best Local Similarity 66.7%; Pred. No. 7.8e+05;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 QLSFXFTK 10  
 |||||  
 Db 1 QSSSFIFHK 9

RESULT 9

AAAY62907  
 ID AAY62907 standard; Peptide; 9 AA.

XX AC AAY62907;

XX DT 02-MAR-2000 (first entry)

XX DE PB-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:2463.

XX KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.

XX OS Synthetic.  
 OS Homo sapiens.

XX PN Key Location/Qualifiers

XX FT Modified-site 1..9

XX FT /note= "the terminal residues are condensed with each  
 other to form a cyclic peptide"

XX FT WO9957149-A2.

XX PN 11-NOV-1999.

XX PF 05-MAY-1999; 99WO-CA00363.

XX PR 05-MAY-1998; 98US-0073040.

XX PR 06-NOV-1998; 98US-0187859.

XX PR 20-JAN-1999; 99US-0234395.

XX PR 08-MAR-1999; 99US-0264516.

XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.

XX PI Blaschuk OW, Gour BJ, Byers S;

XX WPI; 2000-038791/03.

PT New cadherin modulating agents, used for modulating nonclassical  
 PT cadherin-mediated functions for treating e.g. cancers, obesity,  
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
 PT disease -  
 XX  
 PS Claim 72; Page 195; 252pp; English.  
 XX  
 CC The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound  
 CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue  
 CC in a mammal. They can also be used for treating e.g. psoriasis,  
 CC arthritis, age-related macular degeneration, multiple sclerosis and  
 CC diabetes. The products can also be used for detection and diagnosis and  
 CC in bioreactors. AAY60592 to AAY64572 represent specifically claimed  
 CC peptides, and AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent  
 CC sequences used in the exemplification of the present invention.  
 XX  
 SQ Sequence 9 AA;  
 Query Match 37.9%; Score 25; DB 21; Length 9;  
 Best Local Similarity 80.0%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 10 KFDLD 14  
 Db |||:|  
 1 KFDID 5  
 RESULT 10  
 AAY63027  
 ID AAY63027 standard; Peptide; 9 AA.  
 XX  
 AC AAY63027;  
 XX  
 DT 02-MAR-2000 (first entry)  
 XX  
 DE PB-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:2583.  
 XX  
 KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1..9  
 FT /note= "the terminal residues are condensed with each  
 FT other to form a cyclic peptide"  
 XX  
 PN W09957149-A2.  
 XX  
 PD 11-NOV-1999.  
 XX

PF 05-MAY-1999; 99WO-CA00363.  
 XX  
 PR 05-MAY-1998; 98US-0073040.  
 PR 06-NOV-1998; 98US-0187859.  
 PR 20-JAN-1999; 99US-0234395.  
 PR 08-MAR-1999; 99US-0264516.  
 XX  
 PA (ADHE-) ADHEREX TECHNOLOGIES INC.  
 XX  
 PI Blaschuk OW, Gour BJ, Byers S;  
 XX  
 DR WPI: 2000-038791/03.  
 XX  
 PT New cadherin modulating agents, used for modulating nonclassical  
 PT cadherin-mediated functions for treating e.g. cancers, obesity,  
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
 PT disease -  
 XX  
 XX Claim 72; Page 197; 252pp; English.  
 XX  
 CC The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound  
 CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue  
 CC in a mammal. They can also be used for treating e.g. psoriasis, and  
 CC diabetes. The products can also be used for detection and diagnosis and  
 CC in bioreactors. AAY60592 to AAY64572 represent specifically claimed  
 CC peptides, and AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent  
 CC sequences used in the exemplification of the present invention.  
 XX  
 SQ Sequence 9 AA;  
 Query Match 37.9%; Score 25; DB 21; Length 9;  
 Best Local Similarity 80.0%; Pred. No. 7.8e+05;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 10 KFDLD 14  
 Db |||:|  
 1 KFDID 5  
 RESULT 11  
 AAB98514  
 ID AAB98514 standard; peptide; 9 AA.  
 XX  
 AC AAB98514;  
 XX  
 DT 03-AUG-2001 (first entry)  
 XX  
 DE Human TAGD-15 peptide fragment #6.  
 XX  
 KW Human; TAGD-15; cytostatic; vaccine; ovarian tumour; cancer;  
 KW tumour antigen-derived gene 15; serine protease.  
 XX  
 OS Homo sapiens.  
 OS W0200129056-A1.  
 XX  
 PD 26-APR-2001.  
 XX



XX 20-OCT-2000; 2000WO-US29095.  
 XX 20-OCT-1999; 99US-0421213.  
 XX (UYAR-) UNIV ARKANSAS.  
 XX O'Brien TJ, Tanimoto H;  
 XX WPI; 2001-381031/40.  
 XX Novel extracellular serine protease, termed tumor antigen-derived gene  
 PT 15 protein overexpressed in carcinomas and DNA encoding it, for  
 PT diagnosis, treatment, prevention of cancer, particularly breast,  
 PT ovarian cancer -  
 XX Example 9; Page 43; 130pp; English.  
 XX The present invention relates to human tumour antigen-derived gene 15  
 CC (TADG-15) protein and coding sequence (see AAH23601 and AAB98500).  
 CC TADG-15 is an extracellular serine protease. It was found that TADG-15 is  
 CC over-expressed in ovarian tumours. TADG-15 protein or its fragments of  
 CC 9-20 residues that lack TADG-15 protease activity are useful for  
 CC vaccinating an individual against TADG-15, having, suspected of having or  
 CC at risk of getting cancer. The present sequence is one such peptide  
 CC fragment of TADG-15.  
 XX Sequence 9 AA;  
 SQ  
 Query Match 37.9%; Score 25; DB 22; Length 9;  
 Best Local Similarity 55.6%; Pred. No. 7.8e+05;  
 Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 5 SFXTKFDL 13  
 Db 1 SLTFRSFDL 9  
 RESULT 12  
 AAB98560  
 ID AAB98560 standard; peptide; 9 AA.  
 AC AAB98560;  
 XX 03-AUG-2001 (first entry)  
 DT Human TADG-15 peptide fragment #52.  
 XX Human; TADG-15; cytostatic; vaccine; ovarian tumour; cancer;  
 KW tumour antigen-derived gene 15; serine protease.  
 XX Homo sapiens.  
 OS WO200129056-A1.  
 PN 26-APR-2001.  
 XX 20-OCT-2000; 2000WO-US29095.  
 PF 20-OCT-1999; 99US-0421213.  
 XX (UYAR-) UNIV ARKANSAS.  
 PA O'Brien TJ, Tanimoto H;  
 XX WPI; 2001-381031/40.  
 XX Novel extracellular serine protease, termed tumor antigen-derived gene  
 PT 15 protein overexpressed in carcinomas and DNA encoding it, for  
 PT diagnosis, treatment, prevention of cancer, particularly breast,  
 PT ovarian cancer -  
 XX Example 9; Page 44; 130pp; English.

XX The present invention relates to human tumour antigen-derived gene 15  
 CC (TADG-15) protein and coding sequence (see AAH23601 and AAB98500).  
 CC TADG-15 is an extracellular serine protease. It was found that TADG-15 is  
 CC over-expressed in ovarian tumours. TADG-15 protein or its fragments of  
 CC 9-20 residues that lack TADG-15 protease activity are useful for  
 CC vaccinating an individual against TADG-15, having, suspected of having or  
 CC at risk of getting cancer. The present sequence is one such peptide  
 CC fragment of TADG-15.  
 XX Sequence 9 AA;  
 SQ  
 Query Match 37.9%; Score 25; DB 22; Length 9;  
 Best Local Similarity 55.6%; Pred. No. 7.8e+05;  
 Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 5 SFXTKFDL 13  
 Db 1 SLTFRSFDL 9  
 RESULT 13  
 AAY78322  
 ID AAY78322 standard; Protein; 10 AA.  
 XX AC AAY78322;  
 XX 04-MAY-2000 (first entry)  
 DT Anti-zeta-chain antibody 2-B-5 VH-region CDR1 protein sequence.  
 DE Anti-zeta-chain antibody 2-B-5; rat; human zeta chain; immunisation;  
 XX complementary determining region; CDR; autoimmune disease; cytostatic;  
 KW immune deficiency; T-cell malignancy; infectious disease; antiviral;  
 KW immunosuppressive; antimicrobial; immune response modulator; NK-cell.  
 XX Rattus norvegicus.  
 OS WO200003016-A1.  
 PN 20-JAN-2000.  
 XX 09-JUL-1999; 99WO-EP04838.  
 PF 10-JUL-1998; 98EP-0112867.  
 PR (CONN-) CONNEX GMBH.  
 XX Reiter C;  
 PI WPI; 2000-160926/14.  
 DR N-PSDB; AAZ88320.  
 XX New oligonucleotide, polypeptide, antibody useful for treating  
 PT autoimmune disease, immune deficiencies, T-cell malignancies and  
 PT infectious diseases -  
 XX Claim 10; Fig 6; 79pp; English.  
 XX The present invention describes a nucleic acid molecule (I) encoding at  
 CC least one complementary determining region (CDR) of a variable region of  
 CC an antibody which specifically interacts with the extracellular domain of  
 CC the human zeta-chain. The antibody whose CDR of a variable region is  
 CC encoded by (I), is obtained by immunising a rat with Jurkat cells and  
 CC subsequently with a conjugate comprising a carrier molecule and a  
 CC peptide having 11 N-terminal amino acids of a rat zeta-chain. The  
 CC anti-zeta-chain antibody is useful for the treatment and prevention of  
 CC autoimmune diseases, immune deficiencies, T-cell malignancies,  
 CC infectious diseases and the suppression of immune response preferably in  
 CC order to avoid graft rejection after organ transplantation, malignancies,  
 CC or viral infections. The antibody, and fragments of it, can be useful for  
 CC the enhancement or suppression of NK-cell dependent immunity or for the  
 CC treatment of NK-cell derived malignancies. It can also be useful for the

CC determination of zeta-chain or eta-chain expression on NK-cells,  
 CC T-lymphocytes or their precursors. The present sequence represents the  
 CC CDRL of the VH-region of the anti-zeta-chain antibody 2-B-5, produced  
 CC by rats from the present invention.

XX  
 SQ Sequence 10 AA;  
 Query Match 37.9%; Score 25; DB 21; Length 10;  
 Best Local Similarity 37.5%; Pred. No. 2.4e+02;  
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 FXFTKFDL 13  
 : |||:  
 Db 2 YTFSTYDM 9

RESULT 14  
 AAY62908  
 ID AAY62908 standard; Peptide; 10 AA.

XX AC AAY62908;  
 XX DT 02-MAR-2000 (first entry)  
 XX DE PB-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:2464.

XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin-extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.

XX Synthetic.  
 OS Homo sapiens.

XX Key Location/Qualifiers  
 FH Modified-site 1.10  
 FT /note= "the terminal residues are condensed with each  
 FT other to form a cyclic peptide"

XX PN W0957149-A2.

XX PD 11-NOV-1999.

XX PF 05-MAY-1999; 99WO-CA00363.

XX PR 05-MAY-1998; 98US-0073040.

XX PR 06-NOV-1998; 98US-0187859.

XX PR 20-JAN-1999; 99US-0234395.

XX PR 08-MAR-1999; 99US-0264516.

XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.

XX PI Blaschuk OW, Gour BJ, Byers S;

XX DR WPI; 2000-038791/03.

XX New cadherin modulating agents, used for modulating nonclassical  
 PT cadherin-mediated functions for treating e.g. cancers, obesity,  
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
 PT disease

XX PS Claim 72; Page 195; 252pp; English.

XX The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MA can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,

CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound  
 CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue  
 CC in a mammal. They can also be used for treating e.g. psoriasis,  
 CC arthritis, age-related macular degeneration, multiple sclerosis and  
 CC diabetes. The products can also be used for detection and diagnosis and  
 CC in bioreactors. AAY60592 to AAY64572 represent specifically claimed  
 CC peptides, and AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent  
 CC sequences used in the exemplification of the present invention.

XX SQ Sequence 10 AA;

Query Match 37.9%; Score 25; DB 21; Length 10;  
 Best Local Similarity 80.0%; Pred. No. 2.4e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 KFDLD 14

Db 1 KFDID 5

RESULT 15

AAY63028

ID AAY63028 standard; Peptide; 10 AA.

XX AC AAY63028;

XX DT 02-MAR-2000 (first entry)

XX DE PB-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:2584.

XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;  
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FH Modified-site 1.10

FT /note= "the terminal residues are condensed with each  
 FT other to form a cyclic peptide"

XX PN W0957149-A2.

XX PD 11-NOV-1999.

XX PF 05-MAY-1999; 99WO-CA00363.

XX PR 05-MAY-1998; 98US-0073040.

XX PR 06-NOV-1998; 98US-0187859.

XX PR 20-JAN-1999; 99US-0234395.

XX PR 08-MAR-1999; 99US-0264516.

XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.

XX PI Blaschuk OW, Gour BJ, Byers S;

XX DR WPI; 2000-038791/03.

PT New cadherin modulating agents, used for modulating nonclassical  
 PT cadherin-mediated functions for treating e.g. cancers, obesity,  
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological  
 PT disease

XX Claim 72; Page 197; 252pp; English.

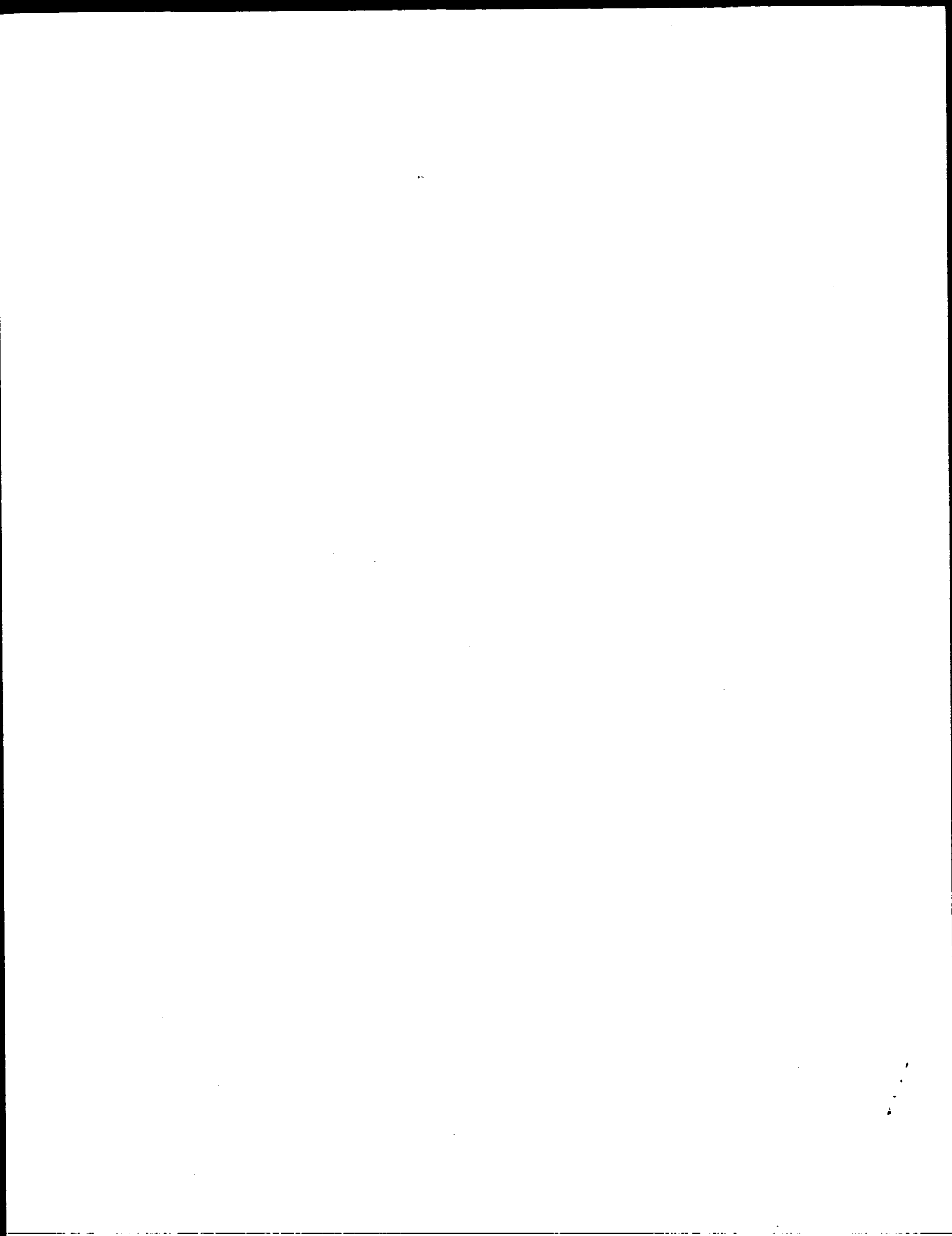
XX The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in  
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
 CC expressing cell, preventing or treating obesity in a mammal, stimulating  
 CC blood vessel regression in a mammal, enhancing drug delivery to the  
 CC central nervous system, treating a demyelinating neurological disease,  
 CC increasing vasopermeability in a mammal, enhancing adhesion of  
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in  
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for  
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound  
 CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue  
 CC in a mammal. They can also be used for treating e.g. psoriasis,  
 CC arthritis, age-related macular degeneration, multiple sclerosis and  
 CC diabetes. The products can also be used for detection and diagnosis and  
 CC in bioreactors. AAY60592 to AAY64572 represent specifically claimed  
 CC peptides, and AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent  
 CC sequences used in the exemplification of the present invention.

XX Sequence 10 AA;

Query Match 37.9%; Score 25; DB 21; Length 10;  
 Best Local Similarity 80.0%; Pred. No. 2.4e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 KFDLD 14  
 |||:  
 Db 1 KFDID 5

Search completed: February 26, 2003, 14:56:22  
 Job time : 35 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:57:21 ; Search time 12 Seconds  
(without alignments)  
44.009 Million cell updates/sec

Title: US-09-476-485A-31  
Perfect score: 66  
Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174566 seqs, 37721826 residues

Total number of hits satisfying chosen parameters: 32941

Minimum DB seq length: 0

Maximum DB seq length: 14

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

PublishedApplications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/FCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/FCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 65    | 98.5        | 14     | 10 | US-09-934-251A-2  |
| 2          | 55    | 83.3        | 12     | 10 | US-09-934-251A-1  |
| 3          | 25    | 37.9        | 13     | 9  | US-09-896-896A-65 |
| 4          | 25    | 37.9        | 13     | 10 | US-09-753-126-101 |
| 5          | 24    | 36.4        | 9      | 10 | US-09-812-079A-7  |
| 6          | 24    | 36.4        | 10     | 9  | US-09-900-590-34  |
| 7          | 24    | 36.4        | 10     | 9  | US-09-791-389-249 |
| 8          | 24    | 36.4        | 10     | 9  | US-09-791-393-249 |
| 9          | 23    | 34.8        | 9      | 10 | US-09-780-053-43  |
| 10         | 23    | 34.8        | 9      | 10 | US-09-780-053-230 |
| 11         | 23    | 34.8        | 9      | 10 | US-09-780-053-317 |
| 12         | 23    | 34.8        | 10     | 10 | US-09-780-053-83  |
| 13         | 23    | 34.8        | 10     | 10 | US-09-780-053-286 |
| 14         | 23    | 34.8        | 10     | 10 | US-09-780-053-391 |
| 15         | 23    | 34.8        | 13     | 9  | US-10-075-846-18  |
| 16         | 23    | 34.8        | 13     | 10 | US-09-870-379-12  |
| 17         | 22    | 33.3        | 14     | 9  | US-09-974-879-475 |
| 18         | 22    | 33.3        | 10     | 9  | US-09-900-590-52  |
| 19         | 22    | 33.3        | 11     | 9  | US-09-921-650-4   |

|    |    |      |    |    |                    |                   |
|----|----|------|----|----|--------------------|-------------------|
| 20 | 22 | 33.3 | 11 | 10 | US-09-874-389-4    | Sequence 4, Appli |
| 21 | 22 | 33.3 | 12 | 10 | US-09-832-723-18   | Sequence 18, Appl |
| 22 | 22 | 33.3 | 13 | 9  | US-09-826-290-359  | Sequence 359, App |
| 23 | 22 | 33.3 | 13 | 9  | US-09-826-290-372  | Sequence 372, App |
| 24 | 22 | 33.3 | 13 | 9  | US-09-791-389-116  | Sequence 116, App |
| 25 | 22 | 33.3 | 13 | 9  | US-09-791-393-116  | Sequence 116, App |
| 26 | 22 | 33.3 | 13 | 10 | US-09-791-378-213  | Sequence 213, App |
| 27 | 22 | 33.3 | 13 | 10 | US-09-791-378-223  | Sequence 223, App |
| 28 | 21 | 31.8 | 9  | 10 | US-10-032-482-14   | Sequence 14, Appl |
| 29 | 21 | 31.8 | 10 | 10 | US-09-780-053-278  | Sequence 278, App |
| 30 | 21 | 31.8 | 10 | 10 | US-09-780-053-709  | Sequence 709, App |
| 31 | 21 | 31.8 | 14 | 9  | US-09-764-868-1242 | Sequence 1242, Ap |
| 32 | 21 | 31.8 | 14 | 9  | US-09-955-999-122  | Sequence 122, App |
| 33 | 20 | 30.3 | 8  | 9  | US-10-079-625-33   | Sequence 33, Appl |
| 34 | 20 | 30.3 | 9  | 8  | US-08-737-457A-8   | Sequence 8, Appli |
| 35 | 20 | 30.3 | 9  | 9  | US-10-106-487-21   | Sequence 21, Appl |
| 36 | 20 | 30.3 | 9  | 9  | US-09-277-074-8    | Sequence 8, Appli |
| 37 | 20 | 30.3 | 9  | 9  | US-09-791-389-166  | Sequence 166, App |
| 38 | 20 | 30.3 | 9  | 9  | US-09-791-393-166  | Sequence 166, App |
| 39 | 20 | 30.3 | 9  | 10 | US-09-759-960-24   | Sequence 24, Appl |
| 40 | 20 | 30.3 | 9  | 10 | US-09-789-720-2    | Sequence 2, Appli |
| 41 | 20 | 30.3 | 9  | 10 | US-09-124-280A-40  | Sequence 40, Appl |
| 42 | 20 | 30.3 | 9  | 10 | US-09-950-313-23   | Sequence 23, Appl |
| 43 | 20 | 30.3 | 9  | 10 | US-09-780-053-619  | Sequence 619, App |
| 44 | 20 | 30.3 | 9  | 10 | US-09-912-787-84   | Sequence 84, Appl |
| 45 | 20 | 30.3 | 9  | 10 | US-09-756-983-5    | Sequence 5, Appli |

#### ALIGNMENTS

##### RESULT 1

US-09-934-251A-2  
; Sequence 2, Application US/09934251A  
; Patent No. US20020132017A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, Jeffrey G.  
; TITLE OF INVENTION: Composition and Method for Preserving Progenitor Cells  
; FILE REFERENCE: 108236.136  
; CURRENT APPLICATION NUMBER: US/09/934, 251A  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 09/368,607  
; PRIOR FILING DATE: 1999-08-05  
; PRIOR APPLICATION NUMBER: US 08/762,537  
; PRIOR FILING DATE: 1996-12-09  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: beta peptide sequence  
US-09-934-251A-2

Query Match 98.5%; Score 65; DB 10; Length 14;  
Best Local Similarity 92.9%; Pred. No. 3.2e-06;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14  
| | | | | | | | | | | | | | | |  
Db 1 AQSLSFXFTKFDLD 14

##### RESULT 2

US-09-934-251A-1  
; Sequence 1, Application US/09934251A  
; Patent No. US20020132017A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, Jeffrey G.  
; TITLE OF INVENTION: Composition and Method for Preserving Progenitor Cells  
; FILE REFERENCE: 108236.136  
; CURRENT APPLICATION NUMBER: US/09/934, 251A

```
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 09/368,607
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 08/762,537
; PRIOR FILING DATE: 1996-12-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide chain of pylartin protein
US-09-934-251A-1

Query Match      83.3%; Score 55; DB 10; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.0002;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXTKFD 12
   ||||| |||||
Db 1 AQSLSFXTKFD 12

RESULT 3
US-09-896-896A-65
; Sequence 65, Application US/09896896A
; Publication No. US20030036181A1
; GENERAL INFORMATION:
; APPLICANT: MAXYGEN APS
; TITLE OF INVENTION: PEPTIDE EXTENDED GLYCOSYLATED POLYPEPTIDES
; FILE REFERENCE: 0217us210
; CURRENT APPLICATION NUMBER: US/09/896,896A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/217,497
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: US 60/225,558
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: DK PA 2000 01027
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DK PA 2000 01092
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: PCT/DK00/00743
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: PCT/DK01/00090
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 65
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-896-896A-65

Query Match      37.9%; Score 25; DB 9; Length 13;
Best Local Similarity 55.6%; Pred. No. 80;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSFXT 9
   ||: |||
Db 1 AQNKTFFNT 9

RESULT 4
US-09-753-126-101
; Sequence 101, Application US/09753126
; Patent No. US20020127219A1
; GENERAL INFORMATION:
; APPLICANT: OKKELS, JENS SIGURD
; APPLICANT: JENSEN, ANNE DAM
```

```
; APPLICANT: HALKIER, TORBEN
; APPLICANT: JENSEN, RIKKE BOLDING
; TITLE OF INVENTION: IMPROVED LYSOSOMAL ENZYMES AND LYSOSOMAL ENZYME
; TITLE OF INVENTION: ACTIVATORS
; FILE REFERENCE: 31-0006000S
; CURRENT APPLICATION NUMBER: US/09/753,126
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: PA 1999 01891
; PRIOR FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 60/174,652
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: PA 200 00865
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/210,984
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 60/211,124
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: PA 2000 01027
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/217,497
; PRIOR FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 101
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-753-126-101

Query Match      37.9%; Score 25; DB 10; Length 13;
Best Local Similarity 55.6%; Pred. No. 80;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSFXT 9
   ||: |||
Db 1 AQNKTFFNT 9

RESULT 5
US-09-812-079A-7
; Sequence 7, Application US/09812079A
; Patent No. US20020058038A1
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles A.
; TITLE OF INVENTION: THERAPEUTIC ANTI-CYTOMEGALOVIRUS
; TITLE OF INVENTION: COMPOUNDS
; FILE REFERENCE: GZ 209500
; CURRENT APPLICATION NUMBER: US/09/812,079A
; CURRENT FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/191,050
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/254,989
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-812-079A-7

Query Match      36.4%; Score 24; DB 10; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 EXFTKF 11
   | | | | |
Db 1 FLFTRF 6
```

RESULT 6  
US-09-900-590-34  
; Sequence 34, Application US/09900590  
; Publication No. US20030028009A1  
; GENERAL INFORMATION:  
; APPLICANT: Huse, William D.  
; Glaser, Scott M.  
; TITLE OF INVENTION: Anti-Alpha V Beta 3 Recombinant Human  
; Antibodies, Nucleic Acids Encoding Same and Methods of Use  
; NUMBER OF SEQUENCES: 100  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell & Flores LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: United States  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/900,590  
; FILING DATE: 06-Jul-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/016,061  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-IX 2965  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 34:  
US-09-900-590-34

Query Match 36.4%; Score 24; DB 9; Length 10;  
Best Local Similarity 37.5%; Pred. No. 94;  
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 FXFTKFDL 13  
| | : | :  
Db 2 FTFSYDM 9

RESULT 7  
US-09-791-389-249  
; Sequence 249, Application US/09791389  
; Publication No. US20030032773A1  
; GENERAL INFORMATION:  
; APPLICANT: Herath, Herath Mudiyansele Athula Chandrasiri  
; APPLICANT: Parekh, Rajesh Bhikhu  
; APPLICANT: Rohlf, Christian  
; APPLICANT: Terrett, Jonathan Alexander  
; APPLICANT: Tyson, Kerry Louise  
; TITLE OF INVENTION: Proteins, Genes and Their Use for  
; TITLE OF INVENTION: Diagnosis and Treatment of Bipolar Affective Disorder (BAD)  
; TITLE OF INVENTION: and Unipolar Depression  
; FILE REFERENCE: 2543-1-001 N2  
; CURRENT APPLICATION NUMBER: US/09/791,389  
; CURRENT FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: GB 0004412.3  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: GB 0030050.9

; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: US 60/254,830  
; PRIOR FILING DATE: 2000-12-12  
; NUMBER OF SEQ ID NOS: 308  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 249  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: homo sapien  
US-09-791-389-249

Query Match 36.4%; Score 24; DB 9; Length 10;  
Best Local Similarity 62.5%; Pred. No. 94;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXFTKFD 12  
| | | | |  
Db 1 SADFTNFD 8

RESULT 8  
US-09-791-393-249  
; Sequence 249, Application US/09791393  
; Publication No. US20030032200A1  
; GENERAL INFORMATION:  
; APPLICANT: Herath, Herath Mudiyansele Athula Chandrasiri  
; APPLICANT: Parekh, Rajesh Bhikhu  
; APPLICANT: Rohlf, Christian  
; TITLE OF INVENTION: Proteins, Genes and Their Use for  
; TITLE OF INVENTION: Diagnosis and Treatment of Bipolar Affective Disorder (BAD)  
; TITLE OF INVENTION: and Unipolar Depression  
; FILE REFERENCE: 2543-1-001 N1  
; CURRENT APPLICATION NUMBER: US/09/791,393  
; CURRENT FILING DATE: 2002-01-02  
; EARLIER APPLICATION NUMBER: GB 0004412.3  
; EARLIER FILING DATE: 2000-02-24  
; EARLIER APPLICATION NUMBER: GB 0030050.9  
; EARLIER FILING DATE: 2000-12-08  
; EARLIER APPLICATION NUMBER: US 60/254,830  
; NUMBER OF SEQ ID NOS: 308  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 249  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: homo sapien  
US-09-791-393-249

Query Match 36.4%; Score 24; DB 9; Length 10;  
Best Local Similarity 62.5%; Pred. No. 94;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXFTKFD 12  
| | | | |  
Db 1 SADFTNFD 8

RESULT 9  
US-09-780-053-43  
; Sequence 43, Application US/09780053  
; Patent No. US20020102640A1  
; GENERAL INFORMATION:  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Mary Faris  
; APPLICANT: Elana Levin  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 83P5G4: A TISSUE SPECIFIC PROTEIN  
; TITLE OF INVENTION: HIGHLY EXPRESSED IN PROSTATE CANCER  
; FILE REFERENCE: 129.505U1  
; CURRENT APPLICATION NUMBER: US/09/780,053

; CURRENT FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/181,261  
; PRIOR FILING DATE: 2000-02-09  
; NUMBER OF SEQ ID NOS: 716  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 43  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-780-053-43

Query Match 34.8%; Score 23; DB 10; Length 9;  
Best Local Similarity 50.0%; Pred. No. 1.5e+05;  
Matches 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 SFXFTKFD 12  
| |:  
Db 1 SVAFSKFE 8

RESULT 10  
US-09-780-053-230  
; Sequence 230, Application US/09780053  
; Patent No. US20020102640A1  
; GENERAL INFORMATION:  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Mary Faris  
; APPLICANT: Elana Levin  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 83P5G4: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.5USU1  
; CURRENT APPLICATION NUMBER: US/09/780,053  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/181,261  
; PRIOR FILING DATE: 2000-02-09  
; NUMBER OF SEQ ID NOS: 716  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 230  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-780-053-230

Query Match 34.8%; Score 23; DB 10; Length 9;  
Best Local Similarity 50.0%; Pred. No. 1.5e+05;  
Matches 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 SFXFTKFD 12  
| |:  
Db 1 SVAFSKFE 8

RESULT 11  
US-09-780-053-317  
; Sequence 317, Application US/09780053  
; Patent No. US20020102640A1  
; GENERAL INFORMATION:  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Mary Faris  
; APPLICANT: Elana Levin  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 83P5G4: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.5USU1  
; CURRENT APPLICATION NUMBER: US/09/780,053  
; CURRENT FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/181,261  
; PRIOR FILING DATE: 2000-02-09  
; NUMBER OF SEQ ID NOS: 716  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 317  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-780-053-317

Query Match 34.8%; Score 23; DB 10; Length 9;  
Best Local Similarity 50.0%; Pred. No. 1.5e+05;  
Matches 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 SFXFTKFD 12  
| |:  
Db 1 SVAFSKFE 8

RESULT 12  
US-09-780-053-83  
; Sequence 83, Application US/09780053  
; Patent No. US20020102640A1  
; GENERAL INFORMATION:  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Mary Faris  
; APPLICANT: Elana Levin  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 83P5G4: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.5USU1  
; CURRENT APPLICATION NUMBER: US/09/780,053  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/181,261  
; PRIOR FILING DATE: 2000-02-09  
; NUMBER OF SEQ ID NOS: 716  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 83  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-780-053-83

Query Match 34.8%; Score 23; DB 10; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 SFXFTKFD 12  
| |:  
Db 2 SVAFSKFE 9

RESULT 13  
US-09-780-053-286  
; Sequence 286, Application US/09780053  
; Patent No. US20020102640A1  
; GENERAL INFORMATION:  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Mary Faris  
; APPLICANT: Elana Levin  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: 83P5G4: A TISSUE SPECIFIC PROTEIN  
; FILE REFERENCE: 129.5USU1  
; CURRENT APPLICATION NUMBER: US/09/780,053  
; CURRENT FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/181,261



; PRIOR FILING DATE: 2000-02-09  
; NUMBER OF SEQ ID NOS: 716  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 286  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-780-053-286

Query Match 34.8%; Score 23; DB 10; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 SFXFTKFD 12  
| | | | |  
Db 2 SVAFSKFE 9

## RESULT 14

US-09-780-053-391  
; Sequence 391, Application US/09780053  
; Patent No. US20020102640A1  
; GENERAL INFORMATION:  
; APPLICANT: Rene S. Hubert  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Mary Faris  
; APPLICANT: Elana Levin  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Ava Jakobovits  
; TITLE OF INVENTION: 83P5G4: A TISSUE SPECIFIC PROTEIN  
; TITLE OF INVENTION: HIGHLY EXPRESSED IN PROSTATE CANCER  
; FILE REFERENCE: 129.5US01  
; CURRENT APPLICATION NUMBER: US/09780,053  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/181,261  
; PRIOR FILING DATE: 2000-02-09  
; NUMBER OF SEQ ID NOS: 716  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 391  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-780-053-391

Query Match 34.8%; Score 23; DB 10; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 SFXFTKFD 12  
| | | | |  
Db 2 SVAFSKFE 9

## RESULT 15

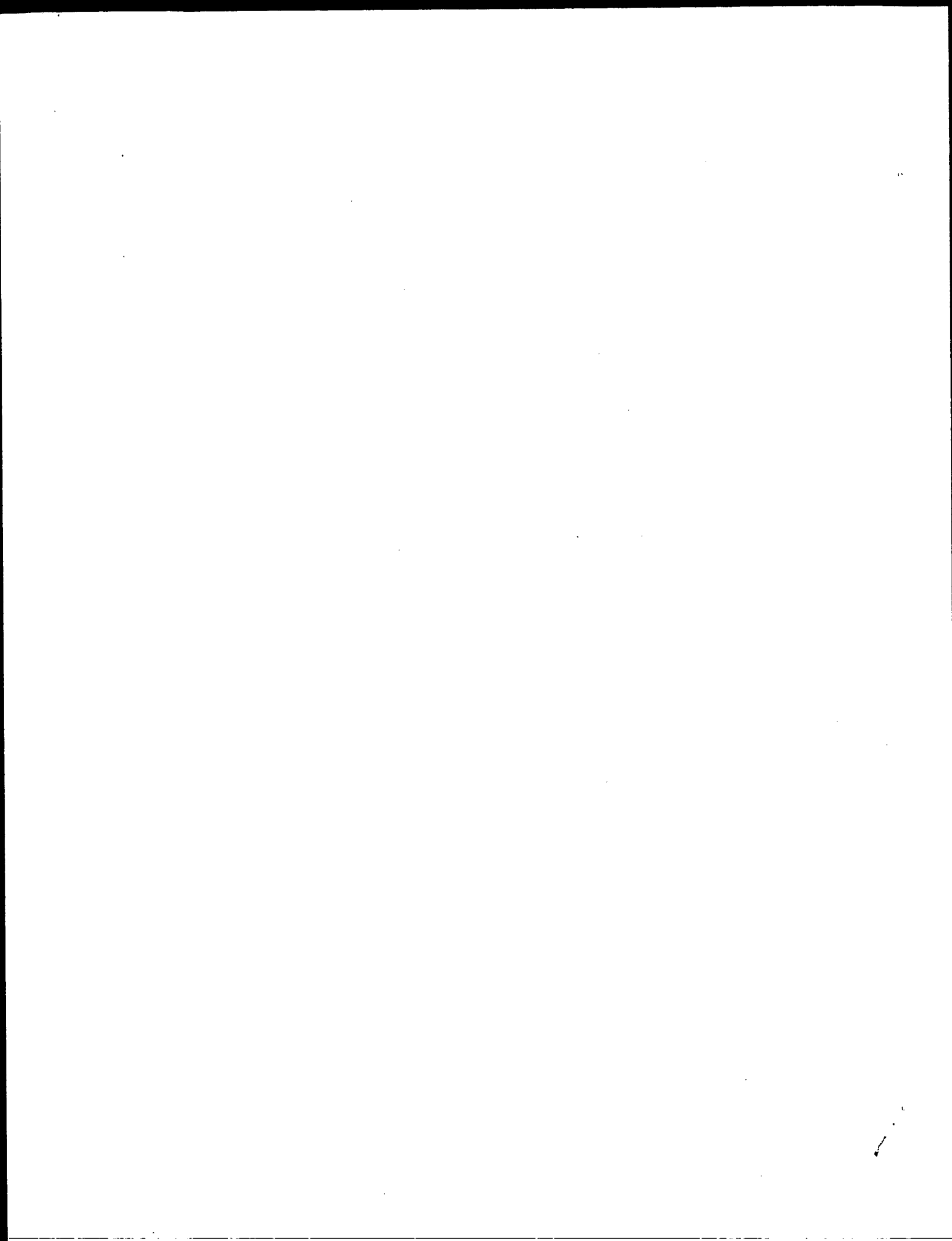
US-10-075-846-18  
; Sequence 18, Application US/10075846  
; Publication No. US20030032608A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT  
; TITLE OF INVENTION: IN THE GASTROINTESTINAL TRACT, HGRA4, and SPLICE VARIANT THEREOF  
; FILE REFERENCE: D0079 NP  
; CURRENT APPLICATION NUMBER: US/10/075,846  
; CURRENT FILING DATE: 2002-02-13  
; PRIOR APPLICATION NUMBER: US 60/269,535  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 18  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: homo sapiens

US-10-075-846-18

Query Match 34.8%; Score 23; DB 9; Length 13;  
Best Local Similarity 55.6%; Pred. No. 1.9e+02;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXFTKFDL 13  
| | | | |  
Db 2 SFGYTMKDL 10

Search completed: February 26, 2003, 15:01:35  
Job time : 13 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:56:26 ; Search time 132 Seconds  
(without alignments)  
68.381 Million cell updates/sec

Title: US-09-476-485A-31

Perfect score: 66

Sequence: 1 AQSLSFXFKFDLD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 509362

Minimum DB seq length: 0

Maximum DB seq length: 14

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

- 1: /cgn2\_6/ptodata/1/paa/PCTUS06\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/paa/US080\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*
- 8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*
- 9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*
- 10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*
- 11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*
- 12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*
- 13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*
- 14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*
- 15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*
- 16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*
- 17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*
- 18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*
- 19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*
- 20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*
- 21: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep.\*
- 22: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*
- 23: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep.\*
- 24: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*
- 25: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*
- 26: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*
- 27: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No.         | Score | Query Match | Length | ID | Description       |
|--------------------|-------|-------------|--------|----|-------------------|
| 1                  | 65    | 98.5        | 14     | 1  | PCT-US97-22486-2  |
| 2                  | 65    | 98.5        | 14     | 23 | US-09-934-251A-2  |
| 3                  | 64    | 97.0        | 14     | 4  | US-08-081-508-12  |
| 4                  | 64    | 97.0        | 14     | 5  | US-08-157-490-11  |
| 5                  | 64    | 97.0        | 14     | 18 | US-09-476-485A-31 |
| 6                  | 55    | 83.3        | 12     | 1  | PCT-US97-22486-1  |
| Sequence 2, Appli  |       |             |        |    |                   |
| Sequence 2, Appli  |       |             |        |    |                   |
| Sequence 12, Appli |       |             |        |    |                   |
| Sequence 11, Appli |       |             |        |    |                   |
| Sequence 31, Appli |       |             |        |    |                   |
| Sequence 1, Appli  |       |             |        |    |                   |

|    |      |      |    |    |                     |                    |
|----|------|------|----|----|---------------------|--------------------|
| 7  | 55   | 83.3 | 12 | 23 | US-09-934-251A-1    | Sequence 1, Appli  |
| 8  | 47   | 71.2 | 14 | 18 | US-09-476-485A-32   | Sequence 32, Appli |
| 9  | 38.5 | 58.3 | 13 | 18 | US-09-476-485A-34   | Sequence 34, Appli |
| 10 | 28   | 42.4 | 12 | 13 | US-08-905-825-67    | Sequence 67, Appli |
| 11 | 27   | 40.9 | 6  | 4  | US-08-080-244-12    | Sequence 12, Appli |
| 12 | 25   | 37.9 | 8  | 22 | US-09-839-542-2462  | Sequence 2462, Ap  |
| 13 | 25   | 37.9 | 8  | 22 | US-09-839-542-2582  | Sequence 2582, Ap  |
| 14 | 25   | 37.9 | 8  | 22 | US-09-839-542B-2462 | Sequence 2462, Ap  |
| 15 | 25   | 37.9 | 8  | 22 | US-09-839-542B-2582 | Sequence 2582, Ap  |
| 16 | 25   | 37.9 | 8  | 24 | US-10-006-869-2462  | Sequence 2462, Ap  |
| 17 | 25   | 37.9 | 8  | 24 | US-10-006-869-2582  | Sequence 2582, Ap  |
| 18 | 25   | 37.9 | 9  | 1  | PCT-US00-29095-24   | Sequence 24, Appli |
| 19 | 25   | 37.9 | 9  | 1  | PCT-US00-29095-70   | Sequence 70, Appli |
| 20 | 25   | 37.9 | 9  | 9  | US-08-589-107-120   | Sequence 120, App  |
| 21 | 25   | 37.9 | 9  | 18 | US-09-421-213-24    | Sequence 24, Appli |
| 22 | 25   | 37.9 | 9  | 18 | US-09-421-213-70    | Sequence 70, Appli |
| 23 | 25   | 37.9 | 9  | 20 | US-09-654-600A-24   | Sequence 24, Appli |
| 24 | 25   | 37.9 | 9  | 20 | US-09-654-600A-70   | Sequence 70, Appli |
| 25 | 25   | 37.9 | 9  | 22 | US-09-839-542-2463  | Sequence 2463, Ap  |
| 26 | 25   | 37.9 | 9  | 22 | US-09-839-542-2583  | Sequence 2583, Ap  |
| 27 | 25   | 37.9 | 9  | 22 | US-09-839-542B-2463 | Sequence 2463, Ap  |
| 28 | 25   | 37.9 | 9  | 22 | US-09-839-542B-2583 | Sequence 2583, Ap  |
| 29 | 25   | 37.9 | 9  | 24 | US-10-006-869-2463  | Sequence 2463, Ap  |
| 30 | 25   | 37.9 | 9  | 24 | US-10-006-869-2583  | Sequence 2583, Ap  |
| 31 | 25   | 37.9 | 10 | 21 | US-09-743-482A-8    | Sequence 8, Appli  |
| 32 | 25   | 37.9 | 10 | 22 | US-09-839-542-2464  | Sequence 2464, Ap  |
| 33 | 25   | 37.9 | 10 | 22 | US-09-839-542-2584  | Sequence 2584, Ap  |
| 34 | 25   | 37.9 | 10 | 22 | US-09-839-542B-2464 | Sequence 2464, Ap  |
| 35 | 25   | 37.9 | 10 | 22 | US-09-839-542B-2584 | Sequence 2584, Ap  |
| 36 | 25   | 37.9 | 10 | 24 | US-10-006-869-2464  | Sequence 2464, Ap  |
| 37 | 25   | 37.9 | 10 | 24 | US-10-006-869-2584  | Sequence 2584, Ap  |
| 38 | 25   | 37.9 | 11 | 23 | US-09-977-827-10    | Sequence 10, Appli |
| 39 | 25   | 37.9 | 12 | 1  | PCT-US00-07946-273  | Sequence 273, App  |
| 40 | 25   | 37.9 | 12 | 19 | US-09-534-717-273   | Sequence 273, App  |
| 41 | 25   | 37.9 | 12 | 19 | US-09-540-018-31    | Sequence 31, Appli |
| 42 | 25   | 37.9 | 12 | 22 | US-09-801-185A-31   | Sequence 31, Appli |
| 43 | 25   | 37.9 | 12 | 22 | US-09-801-185B-31   | Sequence 31, Appli |
| 44 | 25   | 37.9 | 12 | 25 | US-10-133-715-31    | Sequence 31, Appli |
| 45 | 25   | 37.9 | 13 | 21 | US-09-753-126-101   | Sequence 101, App  |

## ALIGNMENTS

RESULT 1  
PCT-US97-22486-2  
; Sequence 2, Application PC/TUS9722486  
; GENERAL INFORMATION:  
; APPLICANT: Moore, Jeffrey G.  
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron, LLP  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11753  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US97/22486  
; FILING DATE: 9-DEC-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/762,537  
; FILING DATE: 9-DEC-1996  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/825,369  
FILING DATE: 28-MAR-1997  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: O'Dea, Sean W.  
REGISTRATION NUMBER: 37690  
REFERENCE/DOCKET NUMBER: 381-21 CIP/PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 822-3550  
TELEFAX: (516) 822-3582  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US97-22486-2

Query Match 98.5%; Score 65; DB 1; Length 14;  
Best Local Similarity 92.9%; Pred. No. 0.0001;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14  
Db 1 AQSLSFSFTKFDLD 14

## RESULT 2

US-09-934-251A-2  
Sequence 2, Application US/09934251A  
GENERAL INFORMATION:  
APPLICANT: Moore, Jeffrey G.  
TITLE OF INVENTION: Composition and Method for Preserving Progenitor Cells  
FILE REFERENCE: 108236.136  
CURRENT APPLICATION NUMBER: US/09/934,251A  
CURRENT FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: US 09/368,607  
PRIOR FILING DATE: 1999-08-05  
PRIOR APPLICATION NUMBER: US 08/762,537  
PRIOR FILING DATE: 1996-12-09  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 14  
TYPE: PPT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: beta peptide sequence  
US-09-934-251A-2

Query Match 98.5%; Score 65; DB 23; Length 14;  
Best Local Similarity 92.9%; Pred. No. 0.0001;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14  
Db 1 AQSLSFSFTKFDLD 14

## RESULT 3

US-08-081-508-12  
Sequence 12, Application US/08081508  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TORIPOINT HEMATOPOIETIC STEM CELL  
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ImClone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.

ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/081,508  
FILING DATE: 19930621  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/975,049  
FILING DATE: 12-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/977,451  
FILING DATE: 19-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/005,941  
FILING DATE: 15-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,272  
FILING DATE: 01-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US UNASSIGNED  
FILING DATE: 09-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US UNASSIGNED  
FILING DATE: 18-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Felt, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-12P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
US-08-081-508-12

Query Match 97.0%; Score 64; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14  
Db 1 AQSLSFXFTKFDLD 14

RESULT 4  
US-08-157-490-11

Sequence 11, Application US/08157490  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ImClone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/157,490  
FILING DATE: 23-SEP-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/975,049  
FILING DATE: 12-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/977,451  
FILING DATE: 19-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/005,941  
FILING DATE: 15-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/045,272  
FILING DATE: 01-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/076022  
FILING DATE: 09-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/080244  
FILING DATE: 18-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/081508  
FILING DATE: 21-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/096759  
FILING DATE: 22-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/125669  
FILING DATE: 23-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-15P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:

LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
US-08-157-490-11

Query Match 97.0%; Score 64; DB 5; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AQSLSFXFTKFDLD 14  
Db 1 AQSLSFXFTKFDLD 14  
|||||

## RESULT 5

US-09-476-485A-31  
Sequence 31, Application US/09476485A  
GENERAL INFORMATION:  
APPLICANT: Colucci, M. Gabriella  
APPLICANT: Chrispeels, Maarten J.  
APPLICANT: Moore, Jeffrey G.  
TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for  
TITLE OF INVENTION: and Products of Their Use  
FILE REFERENCE: 108236.119  
CURRENT APPLICATION NUMBER: US/09/476,485A  
CURRENT FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: US 08/881,189  
PRIOR FILING DATE: 1997-06-24  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 31  
LENGTH: 14  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Peptide corresponding to Pv-FRIL.  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (7)..(7)  
OTHER INFORMATION: Amino acid 7 is Xaa wherein Xaa = Asn, Cys or Ser.  
US-09-476-485A-31

Query Match 97.0%; Score 64; DB 18; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AQSLSFXFTKFDLD 14  
Db 1 AQSLSFXFTKFDLD 14  
|||||

## RESULT 6

PCT-US97-22486-1  
Sequence 1, Application PC/TUS9722486  
GENERAL INFORMATION:  
APPLICANT: Moore, Jeffrey G.  
TITLE OF INVENTION: COMPOSITION AND METHOD FOR  
TITLE OF INVENTION: PRESERVING PROGENITOR CELLS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hoffmann & Baron, LLP  
STREET: 350 Jericho Turnpike  
CITY: Jericho  
STATE: New York  
COUNTRY: USA  
ZIP: 11753  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US97/22486  
; FILING DATE: 9-DEC-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA: 08/762,537  
; FILING DATE: 9-DEC-1996  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/825,369  
; FILING DATE: 28-MAR-1997  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: O'Dea, Sean W.  
; REGISTRATION NUMBER: 37690  
; REFERENCE/DOCKET NUMBER: 381-21 CIP/PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; PCT-US97-22486-1

Query Match 83.3%; Score 55; DB 1; Length 12;  
Best Local Similarity 91.7%; Pred. No. 0.0063;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12  
||||| |||||  
Db 1 AQSLSFSFTKFD 12

RESULT 7  
US-09-934-251A-1  
; Sequence 1, Application US/09934251A  
; GENERAL INFORMATION:  
; APPLICANT: Moore, Jeffrey G.  
; TITLE OF INVENTION: Composition and Method for Preserving Progenitor Cells  
; FILE REFERENCE: 108236.136  
; CURRENT APPLICATION NUMBER: US/09/934,251A  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 09/368,607  
; PRIOR FILING DATE: 1999-08-05  
; PRIOR APPLICATION NUMBER: US 08/762,537  
; PRIOR FILING DATE: 1996-12-09  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: peptide chain of pylartin protein  
US-09-934-251A-1

Query Match 83.3%; Score 55; DB 23; Length 12;  
Best Local Similarity 91.7%; Pred. No. 0.0063;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12  
||||| |||||  
Db 1 AQSLSFSFTKFD 12

RESULT 8  
US-09-476-485A-32

; Sequence 32, Application US/09476485A  
; GENERAL INFORMATION:  
; APPLICANT: Colucci, M. Gabriella  
; APPLICANT: Chrispeels, Maarten J.  
; APPLICANT: Moore, Jeffrey G.  
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for  
; FILING DATE: 9-DEC-1997  
; FILE REFERENCE: 108236.119  
; CURRENT APPLICATION NUMBER: US/09/476,485A  
; CURRENT FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 08/881,189  
; PRIOR FILING DATE: 1997-06-24  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 32  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polypeptide of 18 kDa.  
; NAME/KEY: PEPTIDE  
; LOCATION: (7)..(7)  
; OTHER INFORMATION: Amino acid 7 is Xaa wherein Xaa = unknown amino acid.  
US-09-476-485A-32

Query Match 71.2%; Score 47; DB 18; Length 14;  
Best Local Similarity 85.7%; Pred. No. 0.22;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14  
||||| ||||| ||  
Db 1 AQSLSFXFTKDALD 14

RESULT 9  
US-09-476-485A-34  
; Sequence 34, Application US/09476485A  
; GENERAL INFORMATION:  
; APPLICANT: Colucci, M. Gabriella  
; APPLICANT: Chrispeels, Maarten J.  
; APPLICANT: Moore, Jeffrey G.  
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for  
; FILING DATE: 9-DEC-1997  
; FILE REFERENCE: 108236.119  
; CURRENT APPLICATION NUMBER: US/09/476,485A  
; CURRENT FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 08/881,189  
; PRIOR FILING DATE: 1997-06-24  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 34  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Aminoterminal polypeptide.  
; NAME/KEY: PEPTIDE  
; LOCATION: (7)..(7)  
; OTHER INFORMATION: Amino acid 7 is Xaa wherein Xaa = unknown amino acid.  
US-09-476-485A-34

Query Match 58.3%; Score 38.5; DB 18; Length 13;  
Best Local Similarity 91.7%; Pred. No. 7.7;  
Matches 11; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AQSLSFXFTKFD 12  
||||| ||||| ||  
Db 1 AQSLSFXF-KFD 11

RESULT 10

US-08-905-825-67  
; Sequence 67, Application US/08905825  
; GENERAL INFORMATION:  
; APPLICANT: Watkins, Jeffrey D.  
; APPLICANT: Huse, William D.  
; APPLICANT: Wu, Heiren  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING LIGAND SPECIFIC  
; BINDING MOLECULES  
; NUMBER OF SEQUENCES: 100  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell & Flores, LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: United States  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/905,825  
; FILING DATE: 04-AUG-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-IX 1570  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619)535-9001  
; TELEFAX: (619)535-8949  
; INFORMATION FOR SEQ ID NO: 67:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-905-825-67

Query Match 42.4%; Score 28; DB 13; Length 12;  
Best Local Similarity 40.0%; Pred. No. 6.2e+02;  
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 SPXETKFDLD 14  
|||:|:  
Db 3 TTFPSNYDLN 12

RESULT 11  
US-08-080-244-12  
; Sequence 12, Application US/08080244  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Imclone Systems Incorporated  
; STREET: 180 Varlick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/080,244  
; FILING DATE: 19930618  
; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/679,666  
; FILING DATE: 02-APR-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/906,397  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/975,049  
; FILING DATE: 12-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/977,451  
; FILING DATE: 19-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/005,941  
; FILING DATE: 15-JAN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/045,272  
; FILING DATE: 01-APR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US UNASSIGNED  
; FILING DATE: 09-JUN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: LEM-3-11P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-645-1405  
; TELEFAX: 212-645-2054  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
US-08-080-244-12

Query Match 40.9%; Score 27; DB 4; Length 6;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQSLSF 6  
|||||  
Db 1 AQSLSF 6

RESULT 12  
US-09-839-542-2462  
; Sequence 2462, Application US/09839542  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407D1  
; CURRENT APPLICATION NUMBER: US/09/839,542  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 2462

; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence  
US-09-839-542-2462

Query Match 37.9%; Score 25; DB 22; Length 8;  
Best Local Similarity 80.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 KFDLD 14  
|||:|  
Db 1 KFDID 5

RESULT 13  
US-09-839-542-2582  
; Sequence 2582, Application US/09839542  
; GENERAL INFORMATION:  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407D1  
; CURRENT APPLICATION NUMBER: US/09/839,542  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2582  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence  
US-09-839-542-2582

Query Match 37.9%; Score 25; DB 22; Length 8;  
Best Local Similarity 80.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 KFDLD 14  
|||:|  
Db 1 KFDID 5

RESULT 14  
US-09-839-542B-2462  
; Sequence 2462, Application US/09839542B  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407D1  
; CURRENT APPLICATION NUMBER: US/09/839,542B  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2462  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence  
US-09-839-542B-2462

Query Match 37.9%; Score 25; DB 22; Length 8;

Best Local Similarity 80.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 KFDLD 14  
|||:|  
Db 1 KFDID 5

RESULT 15  
US-09-839-542B-2582  
; Sequence 2582, Application US/09839542B  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Symonds, James Matthew  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; FILE REFERENCE: 100086.407D1  
; CURRENT APPLICATION NUMBER: US/09/839,542B  
; CURRENT FILING DATE: 2001-04-20  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2582  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence  
US-09-839-542B-2582

Query Match 37.9%; Score 25; DB 22; Length 8;  
Best Local Similarity 80.0%; Pred. No. 4.2e+06;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 KFDLD 14  
|||:|  
Db 1 KFDID 5

Search completed: February 26, 2003, 15:00:47  
Job time : 133 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:56:46 ; Search time 23 seconds  
(without alignments)  
55,633 Million cell updates/sec

Title: US-09-476-485A-31

Perfect score: 66

Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 486122 seqs, 91396495 residues

Total number of hits satisfying chosen parameters: 152502

Minimum DB seq length: 0

Maximum DB seq length: 14

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*

- 1: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description          |
|------------|-------|-------------|--------|-------|----------------------|
| 1          | 26    | 39.4        | 9      | 5     | US-09-641-528-13672  |
| 2          | 26    | 39.4        | 9      | 5     | US-09-641-528-45663  |
| 3          | 26    | 39.4        | 9      | 5     | US-09-641-528A-13672 |
| 4          | 26    | 39.4        | 9      | 5     | US-09-641-528A-45663 |
| 5          | 26    | 39.4        | 10     | 5     | US-09-641-528-21738  |
| 6          | 26    | 39.4        | 10     | 5     | US-09-641-528A-21738 |
| 7          | 26    | 39.4        | 11     | 5     | US-09-641-528-18763  |
| 8          | 26    | 39.4        | 11     | 5     | US-09-641-528A-18763 |
| 9          | 24    | 36.4        | 9      | 5     | US-09-641-528-16119  |
| 10         | 24    | 36.4        | 9      | 5     | US-09-641-528-46197  |
| 11         | 24    | 36.4        | 9      | 5     | US-09-641-528A-16119 |
| 12         | 24    | 36.4        | 9      | 5     | US-09-641-528A-46197 |
| 13         | 24    | 36.4        | 10     | 5     | US-09-641-528-23707  |
| 14         | 24    | 36.4        | 10     | 5     | US-09-641-528A-23707 |
| 15         | 24    | 36.4        | 10     | 6     | US-10-283-231-34     |
| 16         | 24    | 36.4        | 11     | 5     | US-09-641-528-16120  |
| 17         | 24    | 36.4        | 11     | 5     | US-09-641-528-19850  |
| 18         | 24    | 36.4        | 11     | 5     | US-09-641-528A-16120 |
| 19         | 24    | 36.4        | 11     | 5     | US-09-641-528A-19850 |
| 20         | 24    | 36.4        | 13     | 1     | PCT-US02-16164-74    |
| 21         | 24    | 36.4        | 13     | 1     | PCT-US02-16164-137   |
| 22         | 24    | 36.4        | 13     | 1     | PCT-US02-16164-187   |
| 23         | 24    | 36.4        | 13     | 1     | PCT-US02-16164-237   |
| 24         | 23    | 34.8        | 9      | 1     | PCT-US02-31642-330   |
| 25         | 23    | 34.8        | 9      | 6     | US-10-283-903-17     |
| 26         | 23    | 34.8        | 9      | 6     | US-10-283-903-146    |

#### ALIGNMENTS

##### RESULT 1

US-09-641-528-13672  
; Sequence 13672, Application US/09641528  
; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard

; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS  
; FILE REFERENCE: 2060.0100001

; CURRENT APPLICATION NUMBER: US/09/641,528  
; PRIOR FILING DATE: 2000-08-15

; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10

; NUMBER OF SEQ ID NOS: 51504

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13672

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Peptide Derived from Human Papillomavirus

US-09-641-528-13672

Query Match

Best Local Similarity

Matches

39.4%; Score 26; DB 5; Length 9;

66.7%; Pred.No.4e+05;

6; Conservative

0; Mismatches

3; Indels

0; Gaps

0;

Qy

5 SFXFTKFDL 13

1111111

Db

1 SFNDTTFDL 9

RESULT 2

US-09-641-528-45663

; Sequence 45663, Application US/09641528

; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro

; APPLICANT: Sidney, John

; APPLICANT: Southwood, Scott

; APPLICANT: Chesnut, Robert

; APPLICANT: Celis, Esteban

; APPLICANT: Grey, Howard

; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS

; FILE REFERENCE: 2060.0100001

Sequence 163, App  
Sequence 233, App  
Sequence 330, App  
Sequence 87, Appl  
Sequence 100, App  
Sequence 387, App  
Sequence 674, App  
Sequence 435, App  
Sequence 3167, App  
Sequence 674, App  
Sequence 21628, A  
Sequence 30698, A  
Sequence 34208, A  
Sequence 41604, A  
Sequence 674, App  
Sequence 21628, A  
Sequence 30698, A  
Sequence 34208, A

9 6 US-10-283-903-163  
9 6 US-10-283-903-233  
9 6 US-10-264-309-330  
10 6 US-10-283-903-87  
10 6 US-10-283-903-100  
10 6 US-10-283-903-387  
10 6 US-10-283-903-674  
12 5 US-09-676-475A-435  
5 1 PCT-US02-36496-3167  
5 6 US-10-293-418-3167  
8 5 US-09-641-528-674  
8 5 US-09-641-528-21628  
8 5 US-09-641-528-30698  
8 5 US-09-641-528-34208  
8 5 US-09-641-528-41604  
8 5 US-09-641-528A-674  
8 5 US-09-641-528A-21628  
8 5 US-09-641-528A-30698  
8 5 US-09-641-528A-34208

27 23 34.8  
28 23 34.8  
29 23 34.8  
30 23 34.8  
31 23 34.8  
32 23 34.8  
33 23 34.8  
34 23 34.8  
35 22 33.3  
36 22 33.3  
37 22 33.3  
38 22 33.3  
39 22 33.3  
40 22 33.3  
41 22 33.3  
42 22 33.3  
43 22 33.3  
44 22 33.3  
45 22 33.3

; CURRENT APPLICATION NUMBER: US/09/641,528  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 45663  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
; US-09-641-528-45663

Query Match 39.4%; Score 26; DB 5; Length 9;  
Best Local Similarity 66.7%; Pred. No. 4e+05;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXTKFDL 13  
|||  
Db 1 SFNDTTFDL 9

RESULT 3  
US-09-641-528A-13672  
; Sequence 13672, Application US/09641528A  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641,528A  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13672  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
; US-09-641-528A-13672

Query Match 39.4%; Score 26; DB 5; Length 9;  
Best Local Similarity 66.7%; Pred. No. 4e+05;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXTKFDL 13  
|||  
Db 1 SFNDTTFDL 9

RESULT 4  
US-09-641-528A-45663  
; Sequence 45663, Application US/09641528A  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641,528A

; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 45663  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
; US-09-641-528A-45663

Query Match 39.4%; Score 26; DB 5; Length 9;  
Best Local Similarity 66.7%; Pred. No. 4e+05;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXTKFDL 13  
|||  
Db 1 SFNDTTFDL 9

RESULT 5  
US-09-641-528-21738  
; Sequence 21738, Application US/09641528  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641,528  
; CURRENT FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: US 60/172,705  
; PRIOR FILING DATE: 1999-12-10  
; NUMBER OF SEQ ID NOS: 51504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 21738  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
; US-09-641-528-21738

Query Match 39.4%; Score 26; DB 5; Length 10;  
Best Local Similarity 66.7%; Pred. No. 28;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXTKFDL 13  
|||  
Db 2 SFNDTTFDL 10

RESULT 6  
US-09-641-528A-21738  
; Sequence 21738, Application US/09641528A  
; GENERAL INFORMATION:  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Sidney, John  
; APPLICANT: Southwood, Scott  
; APPLICANT: Chesnut, Robert  
; APPLICANT: Celis, Esteban  
; APPLICANT: Grey, Howard  
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS  
; FILE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS  
; FILE REFERENCE: 2060.0100001  
; CURRENT APPLICATION NUMBER: US/09/641,528A  
; CURRENT FILING DATE: 2000-08-15

```

; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21738
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528A-21738

```

```

Query Match          39.4%; Score 26; DB 5; Length 10;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 3; Indels 3; Gaps 0;

```

```

QY 5 SFXETKFDL 13
   ||| |||
Db 2 SFNDTTFDL 10

```

## RESULT 7

```

US-09-641-528-18763
; Sequence 18763, Application US/09641528
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; TITLE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS
; FILE REFERENCE: 2060.0100001
; CURRENT APPLICATION NUMBER: US/09/641,528
; CURRENT FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18763
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528-18763

```

```

Query Match          39.4%; Score 26; DB 5; Length 11;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 3; Indels 3; Gaps 0;

```

```

QY 5 SFXETKFDL 13
   ||| |||
Db 3 SFNDTTFDL 11

```

## RESULT 8

```

US-09-641-528A-18763
; Sequence 18763, Application US/09641528A
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; TITLE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS
; FILE REFERENCE: 2060.0100001
; CURRENT APPLICATION NUMBER: US/09/641,528A
; CURRENT FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705

```

```

; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18763
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528A-18763

```

```

Query Match          39.4%; Score 26; DB 5; Length 11;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 3; Indels 3; Gaps 0;

```

```

QY 5 SFXETKFDL 13
   ||| |||
Db 3 SFNDTTFDL 11

```

## RESULT 9

```

US-09-641-528-16119
; Sequence 16119, Application US/09641528
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; TITLE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS
; FILE REFERENCE: 2060.0100001
; CURRENT APPLICATION NUMBER: US/09/641,528
; CURRENT FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16119
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528-16119

```

```

Query Match          36.4%; Score 24; DB 5; Length 9;
Best Local Similarity 44.4%; Pred. No. 4e+05;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 5 SFXETKFDL 13
   ||| |||
Db 1 SFTYKYSL 9

```

## RESULT 10

```

US-09-641-528-46197
; Sequence 46197, Application US/09641528
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; TITLE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS
; FILE REFERENCE: 2060.0100001
; CURRENT APPLICATION NUMBER: US/09/641,528
; CURRENT FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10

```

```

; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46197
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528A-46197

```

```

Query Match          36.4%; Score 24; DB 5; Length 9;
Best Local Similarity 44.4%; Pred. No. 4e+05;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 5 SFXFTKFDL 13
   ||:|:|
DB 1 SFTYPKYSL 9

```

```

RESULT 11
US-09-641-528A-16119
; Sequence 16119, Application US/09641528A
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; FILE REFERENCE: 2060.0100001
; CURRENT APPLICATION NUMBER: US/09/641,528A
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16119
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528A-16119

```

```

Query Match          36.4%; Score 24; DB 5; Length 9;
Best Local Similarity 44.4%; Pred. No. 4e+05;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 5 SFXFTKFDL 13
   ||:|:|
DB 1 SFTYPKYSL 9

```

```

RESULT 12
US-09-641-528A-46197
; Sequence 46197, Application US/09641528A
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; FILE REFERENCE: 2060.0100001
; CURRENT APPLICATION NUMBER: US/09/641,528A
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46197
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528A-46197

```

```

Query Match          36.4%; Score 24; DB 5; Length 9;
Best Local Similarity 44.4%; Pred. No. 4e+05;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 5 SFXFTKFDL 13
   ||:|:|
DB 1 SFTYPKYSL 9

```

```

RESULT 13
US-09-641-528-23707
; Sequence 23707, Application US/09641528
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; FILE REFERENCE: 2060.0100001
; CURRENT APPLICATION NUMBER: US/09/641,528
; CURRENT FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23707
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528-23707

```

```

Query Match          36.4%; Score 24; DB 5; Length 10;
Best Local Similarity 44.4%; Pred. No. 72;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 5 SFXFTKFDL 13
   ||:|:|
DB 2 SFTYPKYSL 10

```

```

RESULT 14
US-09-641-528A-23707
; Sequence 23707, Application US/09641528A
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; FILE REFERENCE: 2060.0100001
; CURRENT APPLICATION NUMBER: US/09/641,528A
; CURRENT FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 23707  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus  
US-09-641-528A-23707

Query Match 36.4%; Score 24; DB 5; Length 10;  
Best Local Similarity 44.4%; Pred. No. 72;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

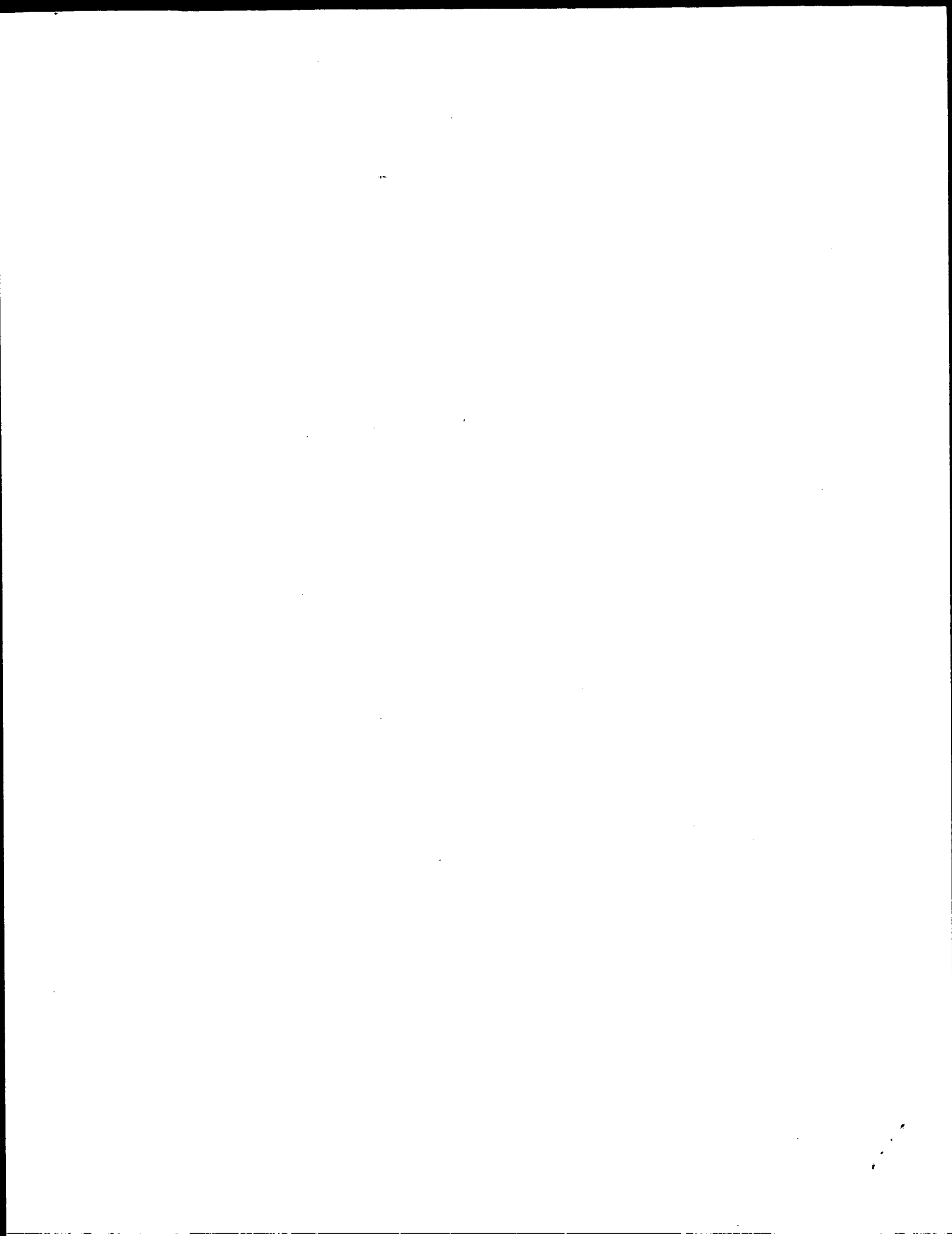
QY 5 SFETKFDL 13  
||:|:|  
Db 2 SFTYPKYSL 10

RESULT 15  
US-10-305-231-34  
; Sequence 34, Application US/10305231  
; GENERAL INFORMATION:  
; APPLICANT: Huse, William D.  
; APPLICANT: Wu, Herren  
; TITLE OF INVENTION: Anti-Alphav Beta3 Recombinant Human Antibodies, Nucleic  
; TITLE OF INVENTION: Acids Encoding Same and Methods of Use  
; FILE REFERENCE: P-IX 3536  
; CURRENT APPLICATION NUMBER: US/10/305,231  
; PRIOR FILING DATE: 2002-11-25  
; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: US/09/339,922  
; PRIOR FILING DATE: CURRENT FILING DATE: 1999-06-24  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-305-231-34

Query Match 36.4%; Score 24; DB 6; Length 10;  
Best Local Similarity 37.5%; Pred. No. 72;  
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 FXETKFDL 13  
||:|:|  
Db 2 FTFSSYDM 9

Search completed: February 26, 2003, 15:01:16  
Job time : 23 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:55:45 : Search time 14 Seconds  
(without alignments)  
29.423 Million cell updates/sec

Title: US-09-476-485A-31

Perfect score: 66

Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 103252

Minimum DB seq length: 0

Maximum DB seq length: 14

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description        |
|------------|-------|---------------|--------|-------|--------------------|
| 1          | 65    | 98.5          | 14     | 3     | US-08-825-369A-2   |
| 2          | 55    | 83.3          | 12     | 3     | US-08-825-369A-1   |
| 3          | 25    | 37.9          | 8      | 4     | US-09-187-859-2462 |
| 4          | 25    | 37.9          | 8      | 4     | US-09-187-859-2582 |
| 5          | 25    | 37.9          | 9      | 4     | US-09-187-859-2463 |
| 6          | 25    | 37.9          | 9      | 4     | US-09-187-859-2583 |
| 7          | 25    | 37.9          | 9      | 4     | US-09-644-600-24   |
| 8          | 25    | 37.9          | 9      | 4     | US-09-644-600-70   |
| 9          | 25    | 37.9          | 10     | 4     | US-09-187-859-2464 |
| 10         | 25    | 37.9          | 10     | 4     | US-09-187-859-2584 |
| 11         | 25    | 37.9          | 11     | 2     | US-08-486-839-9    |
| 12         | 25    | 37.9          | 11     | 3     | US-09-151-011-9    |
| 13         | 25    | 37.9          | 11     | 4     | US-09-343-623-9    |
| 14         | 25    | 37.9          | 12     | 3     | US-08-599-226-31   |
| 15         | 25    | 37.9          | 12     | 4     | US-09-125-098-31   |
| 16         | 24    | 36.4          | 7      | 1     | US-08-092-110A-4   |
| 17         | 24    | 36.4          | 7      | 1     | US-08-273-474-4    |
| 18         | 24    | 36.4          | 7      | 4     | US-08-935-100-4    |
| 19         | 24    | 36.4          | 7      | 5     | PCT-US94-07881-4   |
| 20         | 24    | 36.4          | 8      | 2     | US-08-350-260A-381 |
| 21         | 24    | 36.4          | 10     | 2     | US-08-480-190-149  |
| 22         | 24    | 36.4          | 10     | 2     | US-08-488-379-149  |
| 23         | 24    | 36.4          | 10     | 5     | PCT-US93-07545-149 |
| 24         | 24    | 36.4          | 12     | 1     | US-08-057-184-1    |
| 25         | 24    | 36.4          | 12     | 2     | US-08-480-190-100  |
| 26         | 24    | 36.4          | 12     | 2     | US-08-480-190-111  |
| 27         | 24    | 36.4          | 12     | 2     | US-08-480-190-187  |

|    |    |      |    |   |                    |                   |
|----|----|------|----|---|--------------------|-------------------|
| 28 | 24 | 36.4 | 12 | 2 | US-08-488-379-100  | Sequence 100, App |
| 29 | 24 | 36.4 | 12 | 2 | US-08-488-379-111  | Sequence 111, App |
| 30 | 24 | 36.4 | 12 | 2 | US-08-488-379-187  | Sequence 187, App |
| 31 | 24 | 36.4 | 12 | 5 | PCT-US93-07545-100 | Sequence 100, App |
| 32 | 24 | 36.4 | 12 | 5 | PCT-US93-07545-111 | Sequence 111, App |
| 33 | 24 | 36.4 | 12 | 5 | PCT-US93-07545-187 | Sequence 187, App |
| 34 | 24 | 36.4 | 13 | 1 | US-08-057-184-2    | Sequence 2, Appl  |
| 35 | 24 | 36.4 | 13 | 4 | US-09-171-878-19   | Sequence 19, Appl |
| 36 | 24 | 36.4 | 13 | 4 | US-09-194-285-56   | Sequence 56, Appl |
| 37 | 24 | 36.4 | 14 | 2 | US-08-480-190-110  | Sequence 110, App |
| 38 | 24 | 36.4 | 14 | 2 | US-08-488-379-110  | Sequence 110, App |
| 39 | 24 | 36.4 | 14 | 5 | PCT-US93-07545-110 | Sequence 110, App |
| 40 | 23 | 34.8 | 7  | 4 | US-09-147-933-14   | Sequence 14, Appl |
| 41 | 23 | 34.8 | 8  | 5 | PCT-US95-03236-19  | Sequence 19, Appl |
| 42 | 23 | 34.8 | 12 | 1 | US-08-704-170-54   | Sequence 54, Appl |
| 43 | 23 | 34.8 | 12 | 4 | US-09-258-754-435  | Sequence 435, App |
| 44 | 23 | 34.8 | 12 | 4 | US-09-042-107-435  | Sequence 435, App |
| 45 | 23 | 34.8 | 12 | 5 | PCT-US94-02631-54  | Sequence 54, Appl |

## ALIGNMENTS

RESULT 1  
US-08-825-369A-2  
; Sequence 2, Application US/08825369A  
; Patent No. 6084060  
; GENERAL INFORMATION:  
; APPLICANT: Moore  
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR PRESERVING  
; TITLE OF INVENTION: PROGENITOR CELLS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron, LLP  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11753  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/825,369A  
; FILING DATE: March 28, 1997  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: O'Dea, Sean W.  
; REGISTRATION NUMBER: 37690  
; REFERENCE/DOCKET NUMBER: 381-21 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-825-369A-2

Query Match 98.5%; Score 65; DB 3; Length 14;  
Best Local Similarity 92.9%; Pred. No. 1.9e-06;  
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14  
||||| |||||  
Db 1 AQSLSFXFTKFDLD 14

RESULT 2

US-08-825-369A-1  
; Sequence 1, Application US/08825369A  
; Patent No. 6084060  
; GENERAL INFORMATION:  
; APPLICANT: Moore  
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR PRESERVING  
; TITLE OF INVENTION: PROGENITOR CELLS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann & Baron, LLP  
; STREET: 350 Jericho Turnpike  
; CITY: Jericho  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 11753  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/825,369A  
; FILING DATE: March 28, 1997  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: O'Dea, Sean W.  
; REGISTRATION NUMBER: 37690  
; REFERENCE/DOCKET NUMBER: 381-21 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 822-3550  
; TELEFAX: (516) 822-3582  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-825-369A-1

Query Match 83.3%; Score 55; DB 3; Length 12;  
Best Local Similarity 91.7%; Pred. No. 0.00013;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFSTKFD 12  
| | | | | | | | | |  
DB 1 AQSLSFSTKFD 12

RESULT 3  
US-09-187-859-2462  
; Sequence 2462, Application US/09187859A  
; Patent No. 6358920  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187,859A  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2462  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence  
US-09-187-859-2462

Query Match 37.9%; Score 25; DB 4; Length 8;  
Best Local Similarity 80.0%; Pred. No. 1.9e+05;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 10 KFDLD 14  
| | | | |  
DB 1 KFDID 5  
RESULT 4  
US-09-187-859-2582  
; Sequence 2582, Application US/09187859A  
; Patent No. 6358920  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187,859A  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2582  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence  
US-09-187-859-2582

Query Match 37.9%; Score 25; DB 4; Length 8;  
Best Local Similarity 80.0%; Pred. No. 1.9e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 KFDLD 14  
| | | | |  
DB 1 KFDID 5

RESULT 5  
US-09-187-859-2463  
; Sequence 2463, Application US/09187859A  
; Patent No. 6358920  
; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187,859A  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2463  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence  
US-09-187-859-2463

Query Match 37.9%; Score 25; DB 4; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.9e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 KFDLD 14  
| | | | |  
DB 1 KFDID 5

RESULT 6  
US-09-187-859-2583



; Sequence 2583, Application US/09187859A  
; Patent No. 6358920

; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187,859A  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2583  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence  
US-09-187-859-2583

Query Match 37.9%; Score 25; DB 4; Length 9;  
Best Local Similarity 80.0%; Pred. No. 1.9e+05;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KFDLD 14  
| | | |  
Db 1 KFDID 5

## RESULT 7

US-09-644-600-24  
; Sequence 24, Application US/09644600  
; Patent No. 6451500

; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Tanimoto, Hirotohi  
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease  
; TITLE OF INVENTION: Overexpressed in Carcinomas  
; FILE REFERENCE: D6064CIP/D  
; CURRENT APPLICATION NUMBER: US/09/644,600  
; CURRENT FILING DATE: 2000-08-23  
; PRIOR APPLICATION NUMBER: 09/421,213  
; PRIOR FILING DATE: 1999-10-20  
; PRIOR APPLICATION NUMBER: 09/027,337  
; PRIOR FILING DATE: 1998-02-20  
; NUMBER OF SEQ ID NOS: 98  
; SEQ ID NO 24  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Residues 257-265 of the TADG-15 protein  
US-09-644-600-24

Query Match 37.9%; Score 25; DB 4; Length 9;  
Best Local Similarity 55.6%; Pred. No. 1.9e+05;  
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 SFXFTKFDL 13  
| | | | |  
Db 1 SLTFRSFDL 9

## RESULT 8

US-09-644-600-70  
; Sequence 70, Application US/09644600  
; Patent No. 6451500

; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Tanimoto, Hirotohi  
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease  
; TITLE OF INVENTION: Overexpressed in Carcinomas

; FILE REFERENCE: D6064CIP/D  
; CURRENT APPLICATION NUMBER: US/09/644,600  
; CURRENT FILING DATE: 2000-08-23  
; PRIOR APPLICATION NUMBER: 09/421,213  
; PRIOR FILING DATE: 1999-10-20  
; PRIOR APPLICATION NUMBER: 09/027,337  
; PRIOR FILING DATE: 1998-02-20  
; NUMBER OF SEQ ID NOS: 98  
; SEQ ID NO 70  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Residues 257-265 of the TADG-15 protein  
US-09-644-600-70

Query Match 37.9%; Score 25; DB 4; Length 9;  
Best Local Similarity 55.6%; Pred. No. 1.9e+05;  
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 SFXFTKFDL 13  
| | | | |  
Db 1 SLTFRSFDL 9

## RESULT 9

US-09-187-859-2464  
; Sequence 2464, Application US/09187859A  
; Patent No. 6358920

; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187,859A  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2464  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Representative cyclic modulating agent based on  
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence  
US-09-187-859-2464

Query Match 37.9%; Score 25; DB 4; Length 10;  
Best Local Similarity 80.0%; Pred. No. 57;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KFDLD 14  
| | | | |  
Db 1 KFDID 5

## RESULT 10

US-09-187-859-2584  
; Sequence 2584, Application US/09187859A  
; Patent No. 6358920

; GENERAL INFORMATION:  
; APPLICANT: Blaschuk, Orest W.  
; APPLICANT: Gour, Barbara J.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL  
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS  
; FILE REFERENCE: 100086.407C1  
; CURRENT APPLICATION NUMBER: US/09/187,859A  
; CURRENT FILING DATE: 1998-11-06  
; NUMBER OF SEQ ID NOS: 4052  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2584  
; LENGTH: 10

```

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence
US-09-187-859-2584

Query Match          37.9%; Score 25; DB 4; Length 10;
Best Local Similarity 80.0%; Pred. No. 57;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KFDLD 14
    |||:|
Db 1 KFDID 5

RESULT 11
US-08-486-839-9
; Sequence 9, Application US/08486839
; Patent No. 5928928
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151,011
; FILING DATE: 10 - September - 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Morris, Robert C.
; REGISTRATION NUMBER: 42,910
; REFERENCE/DOCKET NUMBER: 294-32 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-09-151-011-9

Query Match          37.9%; Score 25; DB 3; Length 11;
Best Local Similarity 71.4%; Pred. No. 63;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 8 FTKFDLD 14
    | ||||
Db 1 FDGFDLD 7

RESULT 13
US-09-343-623-9
; Sequence 9, Application US/09343623
; Patent No. 6303118
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/343,623
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:
; TELEFAX:
; INFORMATION FOR SEQ ID NO:
; SEQUENCE CHARACTERISTICS:
; LENGTH:
; TYPE:
; STRANDEDNESS:
; TOPOLOGY:
; MOLECULE TYPE:
; HYPOTHETICAL:

```

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/486,839  
;; FILING DATE: 07-June-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Baron, Ronald J.  
;; REGISTRATION NUMBER: 29,281  
;; REFERENCE/DOCKET NUMBER: 294-26  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (516) 822-3550  
;; TELEFAX: (516) 822-3582  
;; INFORMATION FOR SEQ ID NO: 9:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: unknown  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: protein  
;; HYPOTHETICAL: NO  
;; US-09-343-623-9

Query Match 37.9%; Score 25; DB 4; Length 11;  
Best Local Similarity 71.4%; Pred. No. 63;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 8 FTKFDLD 14  
| | | | |  
Db 1 FDGFDLD 7

## RESULT 14

;; Sequence 31, Application US/0859226  
;; Patent No. 6090382  
;; GENERAL INFORMATION:

;; APPLICANT: Salfeld, Jochen G.  
;; APPLICANT: Allen, Deborah J.  
;; APPLICANT: Hoogenboom, Hendricus R.J.M.  
;; APPLICANT: Kaymakcalan, Zehra  
;; APPLICANT: Labkovsky, Boris  
;; APPLICANT: Mankovich, John A.  
;; APPLICANT: McGuinness, Brian T.  
;; APPLICANT: Roberts, Andrew J.  
;; APPLICANT: Sakorafas, Paul  
;; APPLICANT: Schoenhaut, David  
;; APPLICANT: Vaughan, Tristan J.  
;; APPLICANT: White, Michael  
;; APPLICANT: Wilton, Andrew J.  
;; TITLE OF INVENTION: Human Antibodies that Bind Human TNFa

;; NUMBER OF SEQUENCES: 37  
;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: LAHIVE & COCKFIELD  
;; STREET: 60 State Street, suite 510  
;; CITY: Boston  
;; STATE: Massachusetts  
;; COUNTRY: USA  
;; ZIP: 02109-1875

## COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/599,226  
;; FILING DATE: 08-FEB-1996  
;; CLASSIFICATION: 424

;; ATTORNEY/AGENT INFORMATION:  
;; NAME: DeConti, Giulio A., Jr.  
;; REGISTRATION NUMBER: 31,503  
;; REFERENCE/DOCKET NUMBER: BBI-043  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617)227-7400  
;; TELEFAX: (617)227-5941  
;; INFORMATION FOR SEQ ID NO: 31:

;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 12 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FRAGMENT TYPE: internal  
;; US-08-599-226-31

Query Match 37.9%; Score 25; DB 3; Length 12;  
Best Local Similarity 50.0%; Pred. No. 70;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 5 SFXTKFDLD 14  
| | | | |  
Db 2 SYLSTFSFLD 11

## RESULT 15

;; US-09-125-098-31  
;; Sequence 31, Application US/09125098  
;; Patent No. 6258562  
;; GENERAL INFORMATION:

;; APPLICANT: Salfeld, Jochen G.  
;; APPLICANT: Allen, Deborah J.  
;; APPLICANT: Hoogenboom, Hendricus R.J.M.  
;; APPLICANT: Kaymakcalan, Zehra  
;; APPLICANT: Labkovsky, Boris  
;; APPLICANT: Mankovich, John A.  
;; APPLICANT: McGuinness, Brian T.  
;; APPLICANT: Roberts, Andrew J.  
;; APPLICANT: Sakorafas, Paul  
;; APPLICANT: Schoenhaut, David  
;; APPLICANT: Vaughan, Tristan J.  
;; APPLICANT: White, Michael  
;; APPLICANT: Wilton, Andrew J.  
;; TITLE OF INVENTION: Human Antibodies that Bind Human TNFa

;; NUMBER OF SEQUENCES: 37  
;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: LAHIVE & COCKFIELD  
;; STREET: 60 State Street, suite 510  
;; CITY: Boston  
;; STATE: Massachusetts  
;; COUNTRY: USA  
;; ZIP: 02109-1875

## COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/125,098  
;; FILING DATE:

## CLASSIFICATION:

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/599,226  
;; FILING DATE:

;; ATTORNEY/AGENT INFORMATION:  
;; NAME: DeConti, Giulio A., Jr.  
;; REGISTRATION NUMBER: 31,503  
;; REFERENCE/DOCKET NUMBER: BBI-043  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617)227-7400  
;; TELEFAX: (617)227-5941

;; INFORMATION FOR SEQ ID NO: 31:

;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 12 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FRAGMENT TYPE: internal  
;; US-09-125-098-31

Query Match

37.9%; Score 25; DB 4; Length 12;

Best Local Similarity 50.0%; Pred. No. 70;  
Matches 5; Conservative 1; Mismatches

4; Indels 0; Gaps 0;

Qy 5 SFXTKFDLD 14

1: | | | |

Db 2 SYLSTSFSLD 11

Search completed: February 26, 2003, 14:58:27  
Job time : 14 secs